

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 20, 2005, 01:36:22 ; Search time 7202 Seconds
(without alignments)

11521.814 Million cell updates/sec

Title: US-10-080-522-2

Perfect score: 2180

Sequence: 1 ATCTCTGCTTCTTCTTAGCGT.....TTAAATAAGGTTCTGTCG 2180

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_hic:*

4: gb_est3:*

5: gb_est4:*

6: gb_est5:*

7: gb_est6:*

8: gb_gssi:*

9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1626.6	74.6	1700	3	CR614987
2	1611.6	73.9	1685	3	CR606594
3	1608.6	73.8	1682	3	CR620624
4	1604.6	73.6	1678	3	CR594118
5	1601.6	73.5	1675	3	CR625551
6	1559	71.5	1650	3	CR594943
7	1445.2	66.3	1633	3	CR595956
8	1402.8	64.3	1461	3	CR591775
9	919.6	42.2	1060	5	CR384442
10	916.4	42.0	1047	5	CR379012
11	889.2	40.8	1057	5	CR399171
12	887.2	40.7	1140	5	CR358342
13	882.6	40.5	1026	5	CR417901
14	869.6	39.9	1025	5	CR384441
15	862.8	39.6	1076	5	CR358343
16	861.2	39.5	1011	5	CR396113
17	847.2	38.9	1013	5	CR380489
18	845	38.8	1060	1	AL513766
19	841.4	38.6	1053	5	CR380490
20	824	37.8	1009	5	CR324662
21	822.6	37.7	972	5	CR384475
22	803.6	36.9	1023	5	CR379011
23	791.8	36.3	840	4	BI820744
24	789.4	36.2	1115	4	BM548056

25	776.8	35.6	846	4	BI908417	BI908417	603067363
26	771.6	35.4	1007	5	BM921904	BM921904	AGENCOURT
c	768.4	35.2	1064	5	BM335528	BM335528	AGENCOURT
28	755.6	34.7	920	4	BI821983	BI821983	603040009
29	751.8	34.5	1094	4	BM550009	BM550009	AGENCOURT
30	749.8	34.4	1054	5	BM378128	BM378128	EX378128
31	748.8	34.3	752	4	BI521087	BI521087	603081680
32	746.8	34.3	784	4	BI861812	BI861812	603388835
33	743.4	34.1	850	4	BI763569	BI763569	603050116
34	740.2	34.0	894	4	BI910123	BI910123	603067972
35	737.8	33.8	752	5	CR380144	CR380144	EX380144
36	736.2	33.8	831	5	CR398302	CR398302	EX398302
37	728.4	33.4	959	5	BU158208	BU158208	AGENCOURT
38	727.6	33.4	945	5	CR328120	CR328120	EX328120
39	727.6	33.4	1049	5	CR335529	CR335529	EX335529
40	724.8	33.2	751	4	BI907545	BI907545	603065903
41	722.6	33.1	841	4	BI818929	BI818929	603037421
42	719.6	33.0	824	4	BI909036	BI909036	603070002
c	708	32.5	799	5	CR380143	CR380143	EX380143
44	707.4	32.4	721	4	BI761265	BI761265	603044014
c	704.8	32.3	876	5	EX434442	EX434442	EX434442

ALIGNMENTS

RESULT 1
CR614987
LOCUS
DEFINITION
full-length cDNA clone CSODI025YG13 of Placenta Cot 25-normalized
1700 bp mRNA linear HTC 21-JUL-2004
ACCESSION
CR614987
VERSION
CR614987.1 GI:50495794
KEYWORDS
HTC; CNSLT cDNA.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
REFERENCE
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1700)
AUTHORS
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE
Full-length cDNA libraries and normalization
JOURNAL
Unpublished
REMARK
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
2 (bases 1 to 1700)
REFERENCE
Genoscope.
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequençage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with NotI and cloned
into the NotI and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
Location/Qualifiers
1..1700
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODI025YG13"
/tissue_type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"

ORIGIN
Query Match 74.6%; Score 1626.6; DB 3; Length 1700;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 1688; Conservative 0; Mismatches 9; Indels 6; Gaps 5;
QY 199 TTCTGGGGCTCCGGGGCGCGAGAGTGCATCCAGAGAGCGCGTCCAGAGCGGAC 258
DB 1 TTCTGGGGCTCCGGGGCGCGAGAGTGCATCCAGAGAGCGCGTCCAGAGCGGAC 60

QY	259	CCGGAGTGTTTCAAGAGCCAGTGACAAAGGACGAGGGGCCCAAGTCCCAACGAGCCATGCA	318
DB	61	CCGGAGAGTGTTTCAAGAGCCAGTGACAAAGGACGAGGGGCCCAAGTCCCAACGAGCCATGCA	120
QY	319	GACCTGCCCTCGGCATTCCTGGCCACGCTTTCCAGGCCCTTGGGACCCCTCTGTGTTTT	378
DB	121	GACCTGCCCTCGGCATTCCTGGCCACGCTTTCCAGGCCCTTGGGACCCCTCTGTGTTTT	180
QY	379	GGCTGCCTCTTGGTGTCTCAGAATGAAGGCTGGGACAGCCCATCTGCAAGAGGGGGT	438
DB	181	GGCTGCCTCTTGGTGTCTCAGAATGAAGGCTGGGACAGCCCATCTGCAAGAGGGGGT	240
QY	439	AGTCTCTGTCTTTGGGGCGAGAAACCGGTCAATGCTCTGCAACATCTCCAAGCCCTTCTC	498
DB	241	AGTCTCTGTCTTTGGGGCGAGAAACCGGTCAATGCTCTGCAACATCTCCAAGCCCTTCTC	300
QY	499	CCATGTCAACATCAAGCTGCTGCCCAACGCGGACGAGAGAGGGCATCTTCAATGAGGTGGC	558
DB	301	CCATGTCAACATCAAGCTGCTGCCCAACGCGGACGAGAGAGGGCATCTTCAATGAGGTGGC	360
QY	559	TCGAGGCTACTTCTCCGGGACGGCTGGGAGCTCCAGAGTTCAAGGAGGGCTGGGACAGCT	618
DB	361	TCCAGGCTACTTCTCCGGGACGGCTGGGAGCTCCAGAGTTCAAGGAGGGCTGGGACAGCT	420
QY	619	GGTGATCAAAAGGCGCCCGGACCTCCATGCTGGGCTGTATCATGTGSCA CTTCTGTGGACA	678
DB	421	GGTGATCAAAAGGCGCCCGGACCTCCATGCTGGGCTGTATCATGTGSCA CTTCTGTGGACA	480
QY	679	CCAGAGAAATTAACAGACAGTCAAGCTGGGAGGTTTCAAGTGCAGAACCCAGTTCGCGCCC	738
DB	481	CCAGAGAAATTAACAGACAGTCAAGCTGGGAGGTTTCAAGTGCAGAACCCAGTTCGCGCCC	540
QY	739	TGACACTGGGTTCTTGGGCTGTGCGAGCGGTGGTCACTGCTGTCTTCACTCTCTGTGTCGC	798
DB	541	CGACACTGGGTTCTTGGGCTGTGCGAGCGGTGGTCACTGCTGTCTTCACTCTCTGTGTCGC	600
QY	799	TCTGGTGATGTTGCGCTGGTATCAGGTGCCGTGTTCCAGCAACGCGGGAGNAGAGTT	858
DB	601	TCTGGTGATGTTGCGCTGGTATCAGGTGCCGTGTTCCAGCAACGCGGGAGNAGAGTT	660
QY	859	CTTCTCTCTAGAACCCACAGATGAAGGTCGACGCCCTCAGAGCGGGAGCCAGCAGGGCCCT	918
DB	661	CTTCTCTCTAGAACCCACAGATGAAGGTCGACGCCCTCAGAGCGGGAGCCAGCAGGGCCCT	720
QY	919	GAGCAGAGCCTCGGTGAACCTGTGGA CCCCAGATCCGAGGCCACCCCAAGGCCGTGGC	978
DB	721	GAGCAGAGCCTCGGTGAACCTGTGGA CCCCAGATCCGAGGCCACCCCAAGGCCGTGGC	780
QY	979	ACTGGTGTTCAAA CCCCCTCACACTTGGAGCCCTGGAGCTGCTGTCTCCGCCCAACCCCTTGT	1038
DB	781	ACTGGTGTTCAAA CCCCCTCACACTTGGAGCCCTGGAGCTGCTGTCTCCGCCCAACCCCTTGT	839
QY	1039	TTTCCATATGCCGACAGCCATAGCCGCTTGAAGGCGAGAGGACACAGGAGAGCCAGCC	1098
DB	840	TTTCCATATGCCGACAGCCATAGCCGCTTGAAGGCGAGAGGACACAGGAGAGCCAGCC	899
QY	1099	CTGAGTGCAGACTTGGGTGGGGGCGCTGGGTCTCTGTCCTCCACCCGAGGGGACAGAC	1158
DB	900	CTGAGTGCAGACTTGGGTGGGGGCGCTGGGTCTCTGTCCTCCACCCGAGGGGACAGAC	959
QY	1159	ACCGGCTTGTCTTGGCAGGCTGGGCCCTCTGTGTCA CCACTCTCTGGGTGCTGTGACAGCCCT	1218
DB	960	ACCGGCTTGTCTTGGCAGGCTGGGCCCTCTGTGTCA CCACTCTCTGGGTGCTGTGACAGCCCT	1019
QY	1219	TCCCTTCGACCCCGCAGGCTTTCAGGCTCTGCTTCTCAGTTTCCAAAATGGAAACACC	1278
DB	1020	TCCCTTCGACCCCGCAGGCTTTCAGGCTCTGCTTCTCAGTTTCCAAAATGGAAACACC	1079
QY	1279	TCACTTCCGAGCACCCCGACTTAC CAGGACGATGCCCTCTCCCTCTGCTCATCAAACC	1338
DB	1080	TCACTTCCGAGCACCCCGACTTAC CAGGACGATGCCCTCTCCCTCTGCTCATCAAACC	1139

QY	1339	CACAGACCCGGACTCCCTTTCTGCGACCCAGAGCTGGTCCGGCCCAAGTGTGGGGTCCG	1339
Db	1140	CACAGACCCGGACTCCCTTTCTGCGACCCAGAGCTGGTCCGGCCCAAGTGTGGGGTCCG	1199
QY	1399	CTCTCTCCACTCCACAGGCTCCGGCCCAAGTGAAGGGGGCCCTGCGCGAGCCTCAGACA	1458
Db	1200	CTCTCTCCACTCCACAGGCTCCGGCCCAAGTGAAGGGGGCCCTGCGCGAGCCTCAGACA	1259
QY	1459	CACCTGGAGTTCAAGGC--TGCGGGGGCCCTTGGCAACATACCTGTCTCCCTTGGCTATGAGCAG	1516
Db	1260	CACCTCCAGTTCAAGGCTGTGGGGGGCTTGGCCACATACCTGTCTCCCTTGGCTATGAGCAG	1319
QY	1517	GCTTTGGGGGGCCCTTCCGGCGACGCCGGGGCCGAGGTAGGTCTGCGGGCTTAGAGG	1576
Db	1320	GCTTTGGGGGGCCCTTCCGGCGACGCCGGGGCCGAGGTAGGTCTGCGGGCTTAGAGG	1378
QY	1577	CTGGGATGGCTCTCTGGGCCCAACCGCCAGGGGGCAAGCGCAGCGCGGCTGGAGGCGGGCG	1636
Db	1379	CTGGGATGGCTCTCTGGGCCCAACCGCCAGGGGGCG-AGCGCAGGCGCGGCTGGAGGCGGGCG	1437
QY	1637	GCGCGGGCTCGGGCTGGGGGGTCAAGTGAAGCTG-CCTCGGGGGCTGTCGCGCATCCC	1695
Db	1438	GCGCGGGCTCGGGCTGGGGGGTCAAGTGAAGCTG-CCTCGGGGGCTGTCGCGCATCCC	1497
QY	1696	TCAGTCCCTCGCGCACACCGGGGGTCTGCTCCCTCGTGGCCACCGCACTCGCGAGCCTCTT	1755
Db	1498	TCAGTCCCTCGCGCACACCGGGGGTCTGCTCCCTCGTGGCCACCGCACTCGCGAGCCTCTT	1815
QY	1756	TGGACCCAGATCTGTTATGCTTTTGTCTTCTGTCACCTGCGCGGGGGCCCTTGTATGCTTT	1617
Db	1558	TGGACCCAGATCTGTTATGCTTTTGTCTTCTGTCACCTGCGCGGGGGCCCTTGTATGCTTT	1677
QY	1816	CATCTGTATGGGTGGAAAAATCACCGGAATCCCGGAATCCCGTTCAGTCTCTTGAAAAAGTTCCA	1875
Db	1618	CATCTGTATGGGTGGAAAAATCACCGGAATCCCGGAATCCCGTTCAGTCTCTTGAAAAAGTTCCA	1677
QY	1876	TGACTCGAATATCTGAAATGAAG	1898
Db	1678	TGACTCGAATATCTGAAATGAAG	1700
RESULT 2			
CR606594			
LOCUS			
DEFINITION		1695 bp mRNA linear HTC 21-JUL-2004	
ACCESSION		full-length cDNA clone CS0DI053YJ02 of Placenta Cot 25-normalized	
VERSION		of Homo sapiens (human).	
KEYWORDS		CR606594	
SOURCE		CR606594.1 GI:50487401	
ORGANISM		HTC; CNI5T_cDNA.	
REFERENCE		Homo sapiens (human)	
AUTHORS		Homo sapiens	
TITLE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
JOURNAL		Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
REMARK		1 (bases 1 to 1685)	
		Li, W.B., Gruber, C., Jessee, J. and Polayes, D.	
		Full-length cDNA libraries and normalization	
		Unpublished	
		Contact : Feng Liang Email : fliang@lifetech.com URL :	
		http://fulllength.invitrogen.com/Invitrogen Corporation 1600	
		Faraday Avenue	
		2 (bases 1 to 1685)	
		Genoscope.	
		Direct Submission	
		Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage	
		BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr	
		- Web : www.genoscope.cns.fr)	
		1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prim	
		end enriched, double-strand cDNA was digested with Not I and clone	
		into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library	
		was normalized. Library was constructed by Life Technologies, a	
		division of Invitrogen.	
FEATURES		Location/Qualifiers	
source		1..1685	

http://fulllength.invitrogen.com/ Invitrogen Corporation 1600

Paraday Avenue
2 (bases 1 to 1682)

REFERENCE
AUTHORS
TITLE
JOURNAL

Direct Submission

Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

COMMENT
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.

FEATURES

source

Location/Qualifiers

1..1682
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DK010Y120"
/tissue_type="HeLa cells Cot 25-normalized"
/plasmid="pCMVSPORT_6"

ORIGIN

Query Match 73.8%; Score 1608.6; DB 3; Length 1682;

Best Local Similarity 99.1%; Pred. No. 0;
Matches 1670; Conservative 0; Mismatches 9; Indels 6; Gaps 5;

QY	199	TTCTCGGGCTCGGGGCGGAGAGCTGCATCCAGAGGAGCGCTCCAGGACGGAC	258
DB	1	TTCTCGGGCTCGGGGCGGAGAGCTGCATCCAGAGGAGCGCTCCAGGACGGAC	60
QY	259	CCGGAGAGTGTTCAGAGCCAGTCAAGAGACAGGGGCCAAGTCCACACGACCATGCA	318
DB	61	CCGGAGAGTGTTCAGAGCCAGTGAAGAGACAGGGGCCAAGTCCACACGACCATGCA	120
QY	319	GACCTGCCCCCTGGCATTCCTTGGCCACGTTTCCAGGCCCTTGGAGACCTCCTGTTTT	378
DB	121	GACCTGCCCCCTGGCATTCCTTGGCCACGTTTCCAGGCCCTTGGAGACCTCCTGTTTT	180
QY	379	GGCTGCTCTCTTGAAGTGTCAAGATGAAGGCTGGAGACGCCCATCTGCACAGAGGGGT	438
DB	181	GGCTGCTCTCTTGAAGTGTCAAGATGAAGGCTGGAGACGCCCATCTGCACAGAGGGGT	240
QY	439	AGTCTCTGTGTCTGGGGCGAGAACACCGTCAATGCTTCCAAACATCTCCAAAGCTTCTC	498
DB	241	AGTCTCTGTGTCTGGGGCGAGAACACCGTCAATGCTTCCAAACATCTCCAAAGCTTCTC	300
QY	499	CCATGTCAAACATCAAGCTCGTGCCCAACGGGACAGAGAGCGCCATCTTCAATGAGTGGC	558
DB	301	CCATGTCAAACATCAAGCTCGTGCCCAACGGGACAGAGAGCGCCATCTTCAATGAGTGGC	360
QY	559	TCCAGGCTACTTCTCCGGGACGGCTGGAGCTCCAGGTTCAAGGAGGCGTGGCAGCT	618
DB	361	TCCAGGCTACTTCTCCGGGACGGCTGGAGCTCCAGGTTCAAGGAGGCGTGGCAGCT	420
QY	619	GGTGATCAAAAGGCGCCGGAGCTCCCATGCTGGGCTGTACATGTGGACCTCGTGGACA	678
DB	421	GGTGATCAAAAGGCGCCGGAGCTCCCATGCTGGGCTGTACATGTGGACCTCGTGGACA	480
QY	679	CCAGAGAAATAACAGACAAGTCAAGCTGAGGTTTTCAGGTGAGAAACCCAGTCCGCCCC	738
DB	481	CCAGAGAAATAACAGACAAGTCAAGCTGAGGTTTTCAGGTGAGAAACCCAGTCCGCCCC	540
QY	739	TGACATGGGTTCTGGCTGTGGCAGCGGTGTCACTGTGCTTTCATCTCTTGTGTCG	798
DB	541	CGACACTGGGTTCTGGCTGTGGCAGCGGTGTCACTGTGCTTTCATCTCTTGTGTCG	600
QY	799	TCTGGTCAATGTTCTGGCTGTGGCAGCGGTGTCACTGTGCTTTCATCTCTTGTGTCG	858
DB	601	TCTGGTCAATGTTCTGGCTGTGGCAGCGGTGTCACTGTGCTTTCATCTCTTGTGTCG	660
QY	859	CTTCTCTCTAGAACCCACAGATGAAGTTCGAGCGCCCTCAGAGCGGAGCCGAGGCGCT	918

DB	661	CTTCTCTCTAGAACCCACAGATGAAGTTCGAGCGCCTCAGAGCGGAGCCGAGAGGGGCT	720
QY	919	GAGCAGAGCTCCGCTGAACCTGTGAGCCCGAGAGCTCCGAGCCACCCCAAGCGCGCTGGC	978
DB	721	GAGCAGAGCTCCGCTGAACCTGTGAGCCCGAGAGCTCCGAGCCACCCCAAGCGCGCTGGC	780
QY	979	ACTGGTGTTCAAACCCCTCACCACTTGGAGCCCTGAGTGTCTGTCCTCCCAACCTTGT	1038
DB	781	ACTGGTGTTCAAACCCCTCACCACTTGGAGCCCTGAGTGTCTGTCCTCCCAACCTTGT	839
QY	1039	TTCATATGCCGAGACCCATAGCGCTGCAAGGAGAGAGGACACAGAGAGCCAGCC	1098
DB	840	TTCATATGCCGAGACCCATAGCGCTGCAAGGAGAGAGGACACAGAGAGCCAGCC	899
QY	1099	CTGAGTCCGAGACCTTGGTGGCGGCTGCTCTCTGTCCTCCACCCCGAGAGGACAGAC	1158
DB	900	CTGAGTCCGAGACCTTGGTGGCGGCTGCTCTCTGTCCTCCACCCCGAGAGGACAGAC	959
QY	1159	ACCGGCTTGTGGAGGCTGTGTCACCCACTCTCTGTCACCCACTCTCTGGGTGCTGACGCT	1218
DB	960	ACCGGCTTGTGGAGGCTGTGTCACCCACTCTCTGTCACCCACTCTCTGGGTGCTGACGCT	1019
QY	1219	TCCCTCTCACCCCGGAGCTTTCAGAGCTCTGCTTCTCAGTTTCCAAATGGACACAC	1278
DB	1020	TCCCTCTCACCCCGGAGCTTTCAGAGCTCTGCTTCTCAGTTTCCAAATGGACACAC	1079
QY	1279	TCACCTCCGAGACCCGAGCTTACAGAGACGATGCCCCCTCTCTGCCCCCTCATCAAAAC	1338
DB	1080	TCACCTCCGAGACCCGAGCTTACAGAGACGATGCCCCCTCTCTGCCCCCTCATCAAAAC	1139
QY	1339	CACAGACCCGAGCTTCTGTCACCCAGGCTGTGTCGGGCTCCAGGCTGTGGGCTCG	1398
DB	1140	CACAGACCCGAGCTTCTGTCACCCAGGCTGTGTCGGGCTCCAGGCTGTGGGCTCG	1199
QY	1399	CTCTCTCACTCCAGAGGCTCGGCGCCCAAGTGAAGGGGCCCCCTGCCGAGGCTCAGACA	1458
DB	1200	CTCTCTCACTCCAGAGGCTCGGCGCCCAAGTGAAGGGGCCCCCTGCCGAGGCTCAGACA	1259
QY	1459	CAGTGGAGTTCAGGGCTGGGGGGG--CCTTGCGACATACCTGTCTCTTGGCTATGAGCAG	1516
DB	1260	CAGTGGAGTTCAGGGCTGTGGGGGGCTTGGCCACATACCTGTCTCTTGGCTATGAGCAG	1319
QY	1517	GCTTTGGGGGCTTTCGCGGAGCCCGGGGCGCAGGTAGGGTCTGGGGCTTAGAGG	1576
DB	1320	GCTTTGGGGGCTTTCGCGGAGCCCGGGGCGCAGGTAGGGTCTGGGGCTTAGAGG	1378
QY	1577	CTGGGATGGCTTGGCGCCCAACCGGCGGAGGAGCGAGCGGCGCTGGGAGCGCGG	1636
DB	1379	CTGGGATGGCTTGGCGCCCAACCGGCGGAGGAGCGAGCGGCGCTGGGAGCGCGG	1437
QY	1637	GGCGGGCTCGGGCTGGGGGGTCAAGTGGAGCGCTG--CCTCCGGGGCTGTCTCGCATCC	1695
DB	1438	GGCGGGCTCGGGCTGGGGGGTCAAGTGGAGCGCGGGCTTCCGGGGCTGGAGCGCATCC	1497
QY	1696	TCAGTCCCTTGGGACACCGCGGGGCTGCTCTCTGTCGCCACCGGACCTTCCGAGCTCTT	1755
DB	1498	TCAGTCCCTTGGGACACCGCGGGGCTGCTCTCTGTCGCCACCGGACCTTCCGAGCTCTT	1557
QY	1756	TGGAGCCAGATCTGTTCATGCTTTTGTCTTCTGCTCACTCGCGGGGCCCCCTTTGATGCTT	1815
DB	1558	TGGAGCCAGATCTGTTCATGCTTTTGTCTTCTGCTCACTCGCGGGGCCCCCTTTGATGCTT	1617
QY	1816	CATCTGTATGGGGTGGAAAAATCACCGGGAATCCCCCTTCACTTCTTTGAAAAAGTTCCA	1875
DB	1618	CATCTGTATGGGGTGGAAAAATCACCGGGAATCCCCCTTCACTTCTTTGAAAAAGTTCCA	1677
QY	1876	TGACT 1880	
DB	1678	TGACT 1682	

RESULT 4
CR594118


```

QY 1766 TCTGTTTCATGCTTTTGTCTTCGTCACCTGGCGGGGGCCCTTTTGTCTTTCATCTCTATG 1825
Db 1558 TCTGTTTCATGCTTTTGTCTTCGTCACCTGGCGGGGGCCCTTTTGTCTTTCATCTCTATG 1617
QY 1826 GGTGTGAAATAACACCGGAATCCCTTTCAGTTCTTTTGAATAAGTTTCCATGACTCGAAT 1885
Db 1618 GGTGTGAAATAACACCGGAATCCCTTTCAGTTCTTTTGAATAAGTTTCCATGACTCGAAT 1677

QY 1886 A 1886
Db 1678 A 1678

RESULT 5
LOCUS CR625551 1675 bp mRNA linear HTC 21-JUL-2004
DEFINITION full-length cDNA clone CS0D1029YE07 of Placenta Cot 25-normalized
ACCESSION CR625551
VERSION CR625551.1 GI:50506358
KEYWORDS HTC; CNSLT_cDNA.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1675)
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paraday Avenue
2 (bases 1 to 1675)
REFERENCE Genoscope.
AUTHORS Direct Submission
JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
source Location/Qualifiers
1..1675
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1029YE07"
/tissue_type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"

ORIGIN
Query Match 73.5%; Score 1601.6; DB 3; Length 1675;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 1663; Conservative 0; Mismatches 9; Indels 6; Gaps 5;

QY 210 CCGGGGGCGGAGAGCTGCATCCAGAGAGGGGTCCAGGAGCGGACCGGGAGTGT 269
Db 1 CCGGGGGCGGAGAGCTGCATCCAGAGAGGGGTCCAGGAGCGGACCGGGAGTGT 60

QY 270 TCAAGAGCCAGTGACAAGGACACAGGGGCCCAAGTCCACAGCCATGCAGACTGCCCCC 329
Db 61 TCAAGAGCCAGTGACAAGGACACAGGGGCCCAAGTCCACAGCCATGCAGACTGCCCCC 120

QY 330 TGGCATTCCTGGCCACAGTTTCCAGGCCCTTGGGACCTCTCTGTTTGGTGTGCTCT 389
Db 121 TGGCATTCCTGGCCACAGTTTCCAGGCCCTTGGGACCTCTCTGTTTGGTGTGCTCT 180

QY 390 TGAGTGCTCAGATGAAGCTGGGACAGCCCATCTGACAGAGGGGTAGTCTCTGTGT 449
Db 181 TGAGTGCTCAGATGAAGCTGGGACAGCCCATCTGACAGAGGGGTAGTCTCTGTGT 240

```

```

QY 450 CTTTGGGGGAGAAACACCGTCATGTCCTGCAACATCTCCAACGCCCTTCTCCCATGTCAACA 509
Db 241 CTTTGGGGGAGAAACACCGTCATGTCCTGCAACATCTCCAACGCCCTTCTCCCATGTCAACA 300
QY 510 TCAAGCTGCGTCCCAACCGGGCAGGAGAGCGCCATCTTCAATGAGGTGGCTCCAGGCTACT 569
Db 301 TCAAGCTGCGTCCCAACCGGGCAGGAGAGCGCCATCTTCAATGAGGTGGCTCCAGGCTACT 360
QY 570 TCTCCGGGACCGCTGGCAGCTCCAGGTTTCAGGGAGGGCTGGCAGCAGCTGGTGTATCAAG 629
Db 361 TCTCCGGGACCGCTGGCAGCTCCAGGTTTCAGGGAGGGCTGGCAGCAGCTGGTGTATCAAG 420
QY 630 GGGCCCGGAGCTCCCATGCTGGGGCTGTACATGTGACACCTGTGGGACACCTGTGGGAGAAATA 689
Db 421 GGGCCCGGAGCTCCCATGCTGGGGCTGTACATGTGACACCTGTGGGAGAAATA 480
QY 690 ACAGACAAGTCAACGCTGGAGGTTTTCAGGTGTCAGAAACCCAGTCCCGCCCTGCACACTGGGT 749
Db 481 ACAGACAAGTCAACGCTGGAGGTTTTCAGGTGTCAGAAACCCAGTCCCGCCCTGCACACTGGGT 540
QY 750 TCTGGCCTGTGTCAGCGGTGGTGTCACTGCTGTCTTTCATCTCTTGGTGGTCTGGTGTATGT 809
Db 541 TCTGGCCTGTGTCAGCGGTGGTGTCACTGCTGTCTTTCATCTCTTGGTGGTCTGGTGTATGT 600
QY 810 TCGCCTGTGTACAGGTGCGCTGTTCCAGCAACCGCGGAGAGAAAGTTTCTTCTCTCTAG 869
Db 601 TCGCCTGTGTACAGGTGCGCTGTTCCAGCAACCGCGGAGAGAAAGTTTCTTCTCTCTAG 660
QY 870 AACCCACAGATGAAGTTCGACGCCCTTCAGAGCGGGAGCCAGCAGGGCCCTGCAGAGGCT 929
Db 661 AACCCACAGATGAAGTTCGACGCCCTTCAGAGCGGGAGCCAGCAGGGCCCTGCAGAGGCT 720
QY 930 CGCTGAACTGTGGACCCCGAGACTCCGAGGCCACCCCAAGGCCGCTGGCAGCTGGTGTCA 989
Db 721 CGCTGAACTGTGGACCCCGAGACTCCGAGGCCACCCCAAGGCCGCTGGCAGCTGGTGTCA 780
QY 990 AACCCCTCACCCTTGGAGCCCTGAGGCTGTGTCTCCCGCCCAACCTTGTTCATATGCC 1049
Db 781 AACCCCTCACCCTTGGAGCCCTGAGGCTGTGTGT-CCCCCAACCTTGTTCATATGCC 839
QY 1050 GCAGACCATAGCCCTTCGAAGGAGGAGACACAGGAGAGCCAGCCCTGAGTGCCGA 1109
Db 840 GCAGACCATAGCCCTTCGAAGGAGGAGACACAGGAGAGCCAGCCCTGAGTGCCGA 899
QY 1110 CTTTGGGTGGGGGGCTGGGTCTCTCTGTCACCCACTCTCTGGGTGGGTGAGACACCGGCTTGT 1169
Db 900 CTTTGGGTGGGGGGCTGGGTCTCTCTGTCACCCACTCTCTGGGTGGGTGAGACACCGGCTTGT 959
QY 1170 TGGCAGGCTGGGGCTCTCTGTGTCACCCACTCTCTGGGTGGGTGAGACACCGGCTTGT 1229
Db 960 TGGCAGGCTGGGGCTCTCTGTGTCACCCACTCTCTGGGTGGGTGAGACACCGGCTTGT 1019
QY 1230 CCCCAGGCTTTCAGAGCTCTCTGTGTCACCCACTCTCTGGGTGGGTGAGACACCGGCTTGT 1289
Db 1020 CCCCAGGCTTTCAGAGCTCTCTGTGTCACCCACTCTCTGGGTGGGTGAGACACCGGCTTGT 1079
QY 1290 GCACCCGACTTACAGAGGAGGATGCCCTCTCTGTCACCCACTCTCTGGGTGGGTGAGACACCGG 1349
Db 1080 GCACCCGACTTACAGAGGAGGATGCCCTCTCTGTCACCCACTCTCTGGGTGGGTGAGACACCGG 1139
QY 1350 ACTCTCTTTCGACACCCAGGCTGGTCCGGGCCCGAGGCTGGGTGGGTGAGACACCGGCTTGT 1409
Db 1140 ACTCTCTTTCGACACCCAGGCTGGTCCGGGCCCGAGGCTGGGTGGGTGAGACACCGGCTTGT 1199
QY 1410 CCAGGGCTTCGGGCCCAAGTGAAGGGGCCCTTGCCTGGAGCCCTCAGACACTGAGGTTC 1469
Db 1200 CCAGGGCTTCGGGCCCAAGTGAAGGGGCCCTTGCCTGGAGCCCTCAGACACTTCCAGTTC 1259
QY 1470 AGGGC--TGSGGGGGCTTGGCAGACATCTGTCTCTCTGGCTATGAGCAGGCTTTCGGGGC 1527
Db 1260 AGGGCTTGGGGGGCTTGGGCCACATCTGTCTCTCTGGCTATGAGCAGGCTTTCGGGGC 1319
QY 1528 CTTTCCGGGACGCCCGGGGGCCGAGGTAGGGTCTTGGGGGCTTAGAGGCTGGGATGGCT 1587

```

```
Db      1320 CCTTCGCGGAGCCCCGGGGCCAGGTAGGGTC-GGGGGCTTAGAGGCTGGGATGGCT 1378
Qy      1588 CTGGGCCCCCAGCCAGGCGGCAAGCGCGGCTGGAGGGCGGGCGGGCTCG 1647
Db      1379 CCTGGCCCCCAGCCAGGGGGC-AGGCGAGGCGGGCTGGAGGGCGGGCGGGCTCG 1437
Qy      1648 GGCTGGGGGGTTCAGTGGACGCTG-CCTCGGGGGCTGGTGGGCGATCCTCAGTCCCTCG 1706
Db      1438 GGCTGGGGGGTTCAGTGGAGCGCGGCTCCGGGGCTGGACGGCATCCTCAGTCCCTCG 1497
Qy      1707 GCACCCGGGGTTCCTCGTCGCCACCGCACCTCGCGAGCTCTTTGGACCCAGAT 1766
Db      1498 GCACCCGGGGTTCCTCGTCGCCACCGCACCTCGCGAGCTCTTTGGACCCAGAT 1557
Qy      1767 CTGTTTCATGCTTTTCTTCTGTCCTGCTCACTGGCGGGGCGCTTTGATGCTTTCATCTGATGG 1826
Db      1558 CTGTTTCATGCTTTTCTTCTGTCCTGCTCACTGGCGGGGCGCTTTGATGCTTTCATCTGATGG 1617
Qy      1827 GGTGAAAAAATCACCGGAATCCCTTCAGTTCCTTTGAAAAAGTTCCATGACTCGAA 1884
Db      1618 GGTGAAAAAATCACCGGAATCCCTTCAGTTCCTTTGAAAAAGTTCCATGACTCGAA 1675

RESULT 6
CR594943 1650 bp mRNA linear HTC 21-JUL-2004
LOCUS full-length cDNA clone CS0DE014YC09 of Placenta of Homo sapiens
DEFINITION (human).
ACCESSION CR594943
VERSION CR594943.1 GI:50475750
KEYWORDS HTC; CNSLT_cDNA.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1650)
AUTHORS Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paraday Avenue
Genoscope.
2 (bases 1 to 1650)
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
source
location/Qualifiers
1..1650
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE014YC09"
/tissue_type="Placenta"
/plasmid="pCMVSPORT_6"

ORIGIN
Query Match 71.5%; Score 1559; DB 3; Length 1650;
Best Local Similarity 98.7%; Pred. No. 0;
Matches 1637; Conservative 0; Mismatches 10; Indels 12; Gaps 6;
Qy 258 CCCGGAGTGTTCAGAGCCAGTGCACAGGACCGAGGGGCCCAAGTCCACAGCCATGC 317
Db 1 CCAGGGAGTGTTCAGAGCCAGTGCACAGGACCGAGGGGCCCAAGTCCACAGCCATGC 60
Qy 318 AGACCTGCCCTCGGATTCCTCGCGCAGGTTTCCAGGCGCTTGGGACCTCTCTGTTTT 377
```

```
Db      61 AGACCTGCCCTCGGATTCCTCGCGCAGGTTTCCAGGCGCTTGGGACCTCTCTGTTTT 120
Qy      378 TGGCTGCCCTCTTCTGAGTCTCAGAAATGAAGGCTGGGACAGCCCCCATCTGCAACAGAGGGG 437
Db      121 TGGCTGCCCTCTTCTGAGTCTCAGAAATGAAGGCTGGGACAGCCCCCATCTGCAACAGAGGGG 180
Qy      438 TAGTCTCTGTCTTTGGGGCGAGAACACCGTCATGTCCTGCAACATCTCCAAACGCTTCT 497
Db      181 TAGTCTCTGTCTTTGGGGCGAGAACACCGTCATGTCCTGCAACATCTCCAAACGCTTCT 240
Qy      498 CCCATGTCAAATCAAGCTGGTGGCCACGGGCGAGGAGGCGCATCTTCAATGAGGTGG 557
Db      241 CCCATGTCAAATCAAGCTGGTGGCCACGGGCGAGGAGGCGCATCTTCAATGAGGTGG 300
Qy      558 CTCCAGGCTACTTCTTCCCGGACGCTGGCAGCTCCAGGTTTCAGGAGGCGTGGCACAGC 617
Db      301 CTCCAGGCTACTTCTTCCCGGACGCTGGCAGCTCCAGGTTTCAGGAGGCGTGGCACAGC 360
Qy      618 TGGTGAATCAAAAGGCGCCCGGACCTCCCATGCTGGGCTGTATGTTGGCACTCTGTTGGGAC 677
Db      361 TGGTGAATCAAAAGGCGCCCGGACCTCCCATGCTGGGCTGTATGTTGGCACTCTGTTGGGAC 420
Qy      678 ACCAGAGAAATAACAGACAAGTCAAGCTGGAGGTTTCAGGTCGAGAACCCAGTCCGCC 737
Db      421 ACCAGAGAAATAACAGACAAGTCAAGCTGGAGGTTT-----CAGAACCCAGTCCGCC 474
Qy      738 CTGACACTGGGGTTCCTGGGCTGTGCGAGGGTGGTCACTGCTGTCTTCATCTCTTGGTGG 797
Db      475 CCGACACTGGGGTTCCTGGGCTGTGCGAGGGTGGTCACTGCTGTCTTCATCTCTTGGTGG 534
Qy      798 CTCTGGTTCATGTTCCGCTGGTACAGGTGCGCTGTTCACAGCAACGCGGGGAGAAAGT 857
Db      535 CTCTGGTTCATGTTCCGCTGGTACAGGTGCGCTGTTCACAGCAACGCGGGGAGAAAGT 594
Qy      858 TCTTCTCTAGAACCCAGATGAAGTTCGAGCCCTCAGAGCGGGAGGCCAGAGCGGCC 917
Db      595 TCTTCTCTAGAACCCAGATGAAGTTCGAGCCCTCAGAGCGGGAGGCCAGAGCGGCC 654
Qy      918 TGAGCAGAGCTCCGCTGAACTGTGGACCCAGACTCCGAGCCCAACCAAGGCGCGTGG 977
Db      655 TGAGCAGAGCTCCGCTGAACTGTGGACCCAGACTCCGAGCCCAACCAAGGCGCGTGG 714
Qy      978 CACTGGTGTCAAACCCCTCAACACTTTGAGGCGCTGGAGCTGTGTCCTCCCAACCCCTTG 1037
Db      715 CACTGGTGTCAAACCCCTCAACACTTTGAGGCGCTGGAGCTGTGTCCTCCCAACCCCTTG 773
Qy      1038 TTTCCATATGCGCAGACCCATAGCCGCTTCGAGGAGAGGACAGAGAGCCAGC 1097
Db      774 TTTCCATATGCGCAGACCCATAGCCGCTTCGAGGAGAGGACAGAGAGCCAGC 833
Qy      1098 CTGAGTCCGACCTTTGGGTGGCGGGGCTGGGTCTCTGTCCTCCACCGGAGGCGACAGA 1157
Db      834 CTGAGTCCGACCTTTGGGTGGCGGGGCTGGGTCTCTGTCCTCCACCGGAGGCGACAGA 893
Qy      1158 CACGGCTGTGTTGGCAGGCTGTGTGTCAACCACTCTGGGTGCGTGCAGACCC 1217
Db      894 CACGGCTGTGTTGGCAGGCTGTGTGTCAACCACTCTGGGTGCGTGCAGACCC 953
Qy      1218 TTCCCTCTCCACCCCGGCTCTTCCAGCTCTGCTTCCAGTTTCCAAATGGAACCC 1277
Db      954 TTCCCTCTCCACCCCGGCTCTTCCAGCTCTGCTTCCAGTTTCCAAATGGAACCC 1013
Qy      1278 CTCACCTCCGAGACCCGACTTACCAGAGGCGATGCCCTCTCTGCTCTCATCAAC 1337
Db      1014 CTCACCTCCGAGACCCGACTTACCAGAGGCGATGCCCTCTCTGCTCTCATCAAC 1073
Qy      1338 CCACAGACCCGAGTCTCTTTCTGCAACCCAGGCTGTGCGGCCCGCCAGGTGTGGGGTCC 1397
Db      1074 CCACAGACCCGAGTCTCTCTTTCTGCAACCCAGGCTGTGCGGCCCGCCAGGTGTGGGGTCC 1133
Qy      1398 GCTCTCTCACTCCAGGGCTCCGGGCCCAAGTGAAGGGGGCCCTTCCCGGAGCCTCAGC 1457
```

Db 1134 GCTCTCTCACTCCAGGCGCTCCGGCCCAAGTAGGGGGCCCCCTGCGGAGCCTCAGAC 1193
Qy 1458 AACTGGAGTTAGGCG--TGGGGGGGCTTGGCACAATACCTGTCCCTTGGGTATGACA 1515
Db 1194 AACTCCAGTTAGGCGCTTGGGGGGCTTGGCCACAATACCTGTCCCTTGGGTATGACA 1253
Qy 1516 GGTCTTGGGGGGCTTCCCGCGCAGCCCCGGGGCGAGGTAGGCTTGGGGGCTTAGAG 1575
Db 1254 GGTCTTGGGGGGCTTCCCGCGCAGCCCCGGGGCGAGGTAGGCTTGGGGGCTTAGAG 1312
Qy 1576 GTGGGATGGCTCTTGGCCCCCAGCGGAGGAGCAGCGCGGGCTGGAGAGCGCGC 1635
Db 1313 GTGGGATGGCTCTTGGCCCCCAGCGGAGGAGCAGCGCGGGCTGGAGAGCGCGC 1371
Qy 1636 GGGGGGGCTCGGGCTGGGGGGTCAAGTGGAGCGCTG--CTCGGGGGCTGGTCGCGCATCC 1694
Db 1372 GCGCGGGCTCGGGCTGGGGGGTCAAGTGGAGCGCGCGCTCGGGGGCTGGAGCGCATCC 1431
Qy 1695 CTCAGTCCCTCGGCGCACCGGGGGTGGCTCCCTCGTGGCCACCGCACTGCCGAGCCTCT 1754
Db 1432 CTCAGTCCCTCGGCGCACCGGGGGTGGCTCCCTCGTGGCCACCGCACTGCCGAGCCTCT 1491
Qy 1755 TTGGACCCAGATCTGTTCAATGCTTTTGTCTTCTGTCACCTGCGCGGGGCCCTTTGATGCT 1814
Db 1492 TTGGACCCAGATCTGTTCAATGCTTTTGTCTTCTGTCACCTGCGCGGGGCCCTTTGATGCT 1551
Qy 1815 TCATCTGATGGGTGGAAATATACCGGGAATCCCGCTTCACTTCTTTGAAAGTTCC 1874
Db 1552 TCATCTGATGGGTGGAAATATACCGGGAATCCCGCTTCACTTCTTTGAAAGTTCC 1611
Qy 1875 ATGACTCGAATATCTGAATGAAGAAACAAACCGGACTC 1913
Db 1612 ATGACTCGAATATCTGAATGAAGAAACAAACCGGACTC 1650

RESULT 7

CR595956
LOCUS full-length cDNA clone CS01016YA15 of Placenta Cot 25-normalized
DEFINITION of Homo sapiens (human).
ACCESSION CR595956
VERSION CR595956.1 GI:50476763
KEYWORDS HTC; CNSLT; cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1633)
Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
Contact: Feng Liang Email: fliang@lifetech.com URL:
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
Genoscope.
2 (bases 1 to 1633)
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail: seqref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr)
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
Location/Qualifiers
1. .1633
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS01016YA15"
/tissue_type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"

FEATURES
source

ORIGIN

Query Match 66.3%; Score 1445.2; DB 3; Length 1633;
Best Local Similarity 93.1%; Pred. No. 0;
Matches 1622; Conservative 0; Mismatches 8; Indels 112; Gaps 6;
Qy 199 TTCTTGGGGCTCCGGGGCGCGGAGAGCTGCATCCAGAGGAGCGCTCCAGAGCGGAC 258
Db 1 TTCTTGGGGCTCCGGGGCGCGGAGAGCTGCATCCAGAGGAGCGCTCCAGAGCGGAC 60
Qy 259 CCGGAGAGTGTTCAGAGCCAGTGCAAGAGCAGGGGCCCCAAGTCCCAACAGCCATGCA 318
Db 61 CCGGAGAGTGTTCAGAGCCAGTGCAAGAGCAGGGGCCCCAAGTCCCAACAGCCATGCA 120
Qy 319 GACCTGCCCCCTTGGCAATTCCTTGGCCACGTTTCCAGAGCGCTTGGGAGCCCTCTTTT 378
Db 121 GACCTGCCCCCTTGGCAATTCCTTGGCCACGTTTCCAGAGCGCTTGGGAGCCCTCTTTT 180
Qy 379 GGCTGCTCTTGTAGTGTCTAGATGAAGGCTGGGAGCAGCCCATCTGCACAGAGGGGT 438
Db 181 GGCTGCTCTTGTAGTGTCTAGATGAAGGCTGGGAGCAGCCCATCTGCACAGAGGGGT 240
Qy 439 AGTCTCTGTGCTTGGGCGGAGAACACCGTCAATGCTTGCACAACTCTCCAAAGCCTTCTC 498
Db 241 AGTCTCTGTGCTTGGGCGGAGAACACCGTCAATGCTTGCACAACTCTCCAAAGCCTTCTC 300
Qy 499 CCNTGTCAACATCAGCTGCGTCCCGGAGGAGGAGCGCCATCTTCAATGAGGTGCG 558
Db 301 CCNTGTCAACATCAGCTGCGTCCCGGAGGAGGAGCGCCATCTTCAATGAGGTGCG 360
Qy 559 TCCAGGCTACTTCTCCCGGAGCGCTGGCAGCTCCAGGTTTCAAGGAGCGTGGCAGAGT 618
Db 361 TCCAGGCTACTTCTCCCGGAGCGCTGGCAGCTCCAGGTTTCAAGGAGCGTGGCAGAGT 420
Qy 619 GGTGATCAAGGCGCGCGGAGCTCCCATGCTGGGGTGTACATGTGGGACCTCGTGAGACA 678
Db 421 GGTGATCAAGGCGCGCGGAGCTCCCATGCTGGGGTGTACATGTGGGACCTCGTGAGACA 480
Qy 679 CCAGAGAAATACAGACAAAGTACAGCTCGAGGTTTTCAGGTGCAGAAACCCAGTCCGCC 738
Db 481 CCAGAGAAATACAGACAAAGTACAGCTCGAGGTTTTCAGGTGCAGAAACCCAGTCCGCC 515
Qy 739 TGACACTGGGTTCTGGCCCTGTGCCAGCGGTGGTCACTGCTGTCTTTCATCTCTTGGTGC 798
Db 516 TGACACTGGGTTCTGGCCCTGTGCCAGCGGTGGTCACTGCTGTCTTTCATCTCTTGGTGC 515
Qy 799 TCTGTGTGATTTTGGCTGTGTCAGGTGCGCTGTTCAGAGAACCGCGGAGAGAGTT 858
Db 516 TCTGTGTGATTTTGGCTGTGTCAGGTGCGCTGTTCAGAGAACCGCGGAGAGAGTT 554
Qy 859 CTTCTCTCTAGAACCCCGAGATGAAGTCCAGAGCGCTCAGAGCGGAGCGCCAGAGGGCT 918
Db 555 CTTCTCTCTAGAACCCCGAGATGAAGTCCAGAGCGCTCAGAGCGGAGCGCCAGAGGGCT 614
Qy 919 GAGCAGAGCTCCCGCTGAACTGTGGACCCCGAGACTCCGAGGCCACCCCAAGGCCCTGCG 978
Db 615 GAGCAGAGCTCCCGCTGAACTGTGGACCCCGAGACTCCGAGGCCACCCCAAGGCCCTGCG 674
Qy 979 ACTGTGTTTAAACCTCACAATCTGGAGCGCTTGGAGTGTCTGTCCTCCCTCCCAACCTTGT 1038
Db 675 ACTGTGTTTAAACCTCACCCTTTGGAGCGCTTGGAGTGTCTGTCCTCCCTCCCAACCTTGT 733
Qy 1039 TTCCATATGCGCAGACCCATAGCGCTGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1098
Db 734 TTCCATATGCGCAGACCCATAGCGCTGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 793
Qy 1099 CTGAGTCCGAGCTTGGGTGGCGGGCTGCGGTCTCTCTGTCCTCCAGCGGAGGAGGAGGAG 1158
Db 794 CTGAGTCCGAGCTTGGGTGGCGGGCTGCGGTCTCTCTGTCCTCCAGCGGAGGAGGAGGAG 853
Qy 1159 ACCGGCTTGTGGCAGGCTGGGCTCTGTCACCCACTCTCTGGGTGGTGGTGGTGGTGGT 1218
Db 854 ACCGGCTTGTGGCAGGCTGGGCTCTGTCACCCACTCTCTGGGTGGTGGTGGTGGTGGT 913

```
QY 1219 TCCCTTCGACCCCGAGGTTTCCAGAGCTGTGTTCTCAGTTCAGTTTCCAAATGAAACCACC 1278
|||
Db 914 TCCCTTCGACCCCGAGGTTTCCAGAGCTGTGTTCTCAGTTCAGTTTCCAAATGAAACCACC 973
|||
QY 1279 TCACCTCCGACGACCCGAGTTCACAGGAGCGATGCCCTCCTCTGCTCCCTCATCAAAACC 1338
|||
Db 974 TCACCTCCGACGACCCGAGTTCACAGGAGCGATGCCCTCCTCTGCTCCCTCATCAAAACC 1033
|||
QY 1339 CACAGACCCGAGTCTCCCTTTCTGCCACCCAGGCTGTGTCGGCCCGCAGAGTGTGGGTCGG 1398
|||
Db 1034 CACAGACCCGAGTCTCCCTTTCTGCCACCCAGGCTGTGTCGGCCCGCAGGTTGGGTCGG 1093
|||
QY 1399 CTCTCTCCGACGAGGCTCCGCGCCCAAGTGAAGGGGCCCTGCGGAGGCTCAGACA 1458
|||
Db 1094 CTCTCTCCGACGAGGCTCCGCGCCCAAGTGAAGGGGCCCTGCGGAGGCTCAGACA 1153
|||
QY 1459 CACTGGAGTTCAGGGC--TGGGGGGGCTTGGCCACATCTCTCCCTTGGCTATGAGCAG 1516
|||
Db 1154 CACTCGAGTTCAGGGCTGTGGGGGGGCTTGGCCACATCTCTCTCTGGCTATGAGCAG 1213
|||
QY 1517 GCTTTGGGGGCTCTTCCGCGGCGAGCCCGGGGGCGAGGTAGGGTCTTGGGGCTTAGAGG 1576
|||
Db 1214 GCTTTGGGGGCTCTTCCGCGGCGAGCCCGGGGGCGAGGTAGGGTCTTGGGGCTTAGAGG 1272
|||
QY 1577 CTGGAGTGGCTCTTGGGCCCAACCGCAGGGGCAAGCGCAGGCCCGGGCTTGGAGCGGGCG 1636
|||
Db 1273 CTGGAGTGGCTCTTGGGCCCAACCGCAGGGGCAAGCGCAGGGGCTTGGAGCGGGCG 1331
|||
QY 1637 GGGGGGCTCGGGCTGGGGGTCAGGTGGAAGCTG--CTTCCGGGGCTGTGCGGATCCC 1695
|||
Db 1332 GCGGGGCTCGGGCTGGGGGTCAGGTGGAAGCTG--CTTCCGGGGCTGTGCGGATCCC 1391
|||
QY 1696 TCAGTCCCTCGGCGACCCCGGGGCTGGCTCCCTCGTGGCCCAACCGCAGGCTTGGAGCTCTT 1755
|||
Db 1392 TCAGTCCCTCGGCGACCCCGGGGTCGGTCCCTCGTGGCCCAACCGCAGGCTTGGAGCTCTT 1451
|||
QY 1756 TGAACCCAGATCTGTTCATGCTTTTGTCTTCTGTCATCTGCGGGGGGCTTGTGATGCTCT 1815
|||
Db 1452 TGAACCCAGATCTGTTCATGCTTTTGTCTTCTGTCATCTGCGGGGGGCTTGTGATGCTCT 1511
|||
QY 1816 CATCTGTATGGGTGGAAAAATCACGGGAATCCCCCTTCAGTCTTTTGAAGAAATCCA 1875
|||
Db 1512 CATCTGTATGGGTGGAAAAATCACGGGAATCCCCCTTCAGTCTTTTGAAGAAATCCA 1571
|||
QY 1876 TGACTCGAATATCTGAATGAAGAAAAACAAACCGACTCAAAACCTTCAAGTAGCTCCAA 1935
|||
Db 1572 TGACTCGAATATCTGAATGAAGAAAAACAAACCGACTCAAAACCTTCAAGTAGCTCCAA 1631
|||
QY 1936 AT 1937
|||
Db 1632 AT 1633
```

```
RESULT 8
CR591775
LOCUS full-length cDNA clone CS0DK010YL19 of HeLa cells H7C 21-JUL-2004
DEFINITION of Homo sapiens (human).
ACCESSION CR591775
VERSION CR591775.1 GI:50472582
KEYWORDS H7C; CNSLT cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1461)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
```

```
REFERENCE 2 (bases 1 to 1461)
Genoscope.
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr)
COMMENT: 1st strand cDNA was primed with a NoI-oligo(dT) primer. Five clones
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
Location/Qualifiers
source
1..1461
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DK010YL19"
/tissue_type="HeLa cells Cot 25-normalized"
/plasmid="pCMVSPORT_6"
ORIGIN
Query Match 64.3%; Score 1402.8; DB 3; Length 1461;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1452; Conservative 0; Mismatches 7; Indels 5; Gaps 4;
QY 199 TTCTTGGGGCTCCGGGGCGCGAGAAAGCTTCCATCCAGAGGAGCGCTCCAGAGCGGAC 258
|||
Db 1 TTCTTGGGGCTCCGGGGCGCGAGAAAGCTTCCATCCAGAGGAGCGCTCCAGAGCGGAC 60
|||
QY 259 CCGGAGAGTGTTCAGAGCCAGTGAAGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 318
|||
Db 61 CCGGAGAGTGTTCAGAGCCAGTGAAGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 120
|||
QY 319 GACCTGCCCCCTGGCATTTCCCTGGGCGACGTTTCCAGAGGAGGAGGAGGAGGAGGAGGAG 378
|||
Db 121 GACCTGCCCCCTGGCATTTCCCTGGGCGACGTTTCCAGAGGAGGAGGAGGAGGAGGAGGAG 180
|||
QY 379 GGTGCTCTCTTGTAGTGTCTGAGATGAAGGCTGGGAGAGGAGGAGGAGGAGGAGGAGGAG 438
|||
Db 181 GGTGCTCTCTTGTAGTGTCTGAGATGAAGGCTGGGAGAGGAGGAGGAGGAGGAGGAGGAG 240
|||
QY 439 AGTCTCTGTGTCTGGGGGAGAGACACCGTGTCTGTCATGTCCTGCAACATCTCCAGGCTTCTC 498
|||
Db 241 AGTCTCTGTGTCTGGGGGAGAGACACCGTGTCTGTCATGTCCTGCAACATCTCCAGGCTTCTC 300
|||
QY 499 CCATGTCAACATCAAGCTGCGTGCCACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 558
|||
Db 301 CCATGTCAACATCAAGCTGCGTGCCACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
|||
QY 559 TCCAGGCTACTTCTCCCGGGAGCGGTGGCAGCTCCAGGTTTCAGGAGGAGGAGGAGGAGGAGGAG 618
|||
Db 361 TCCAGGCTACTTCTCCCGGGAGCGGTGGCAGCTCCAGGTTTCAGGAGGAGGAGGAGGAGGAGGAG 420
|||
QY 619 GGTGATCAAGGCGCGCGGAGTCCCATGCTGGGGTGTACATGTGGGAGGAGGAGGAGGAGGAGGAG 678
|||
Db 421 GGTGATCAAGGCGCGCGGAGTCCCATGCTGGGGTGTACATGTGGGAGGAGGAGGAGGAGGAGGAG 480
|||
QY 679 CCAGAGAAATTAACACAGAGTCAAGCTCGAGGTTTCAGGTGCGAGAGGAGGAGGAGGAGGAGGAG 738
|||
Db 481 CCAGAGAAATTAACACAGAGTCAAGCTCGAGGTTTCAGGTGCGAGAGGAGGAGGAGGAGGAGGAG 540
|||
QY 739 TGACACTGGGTTCTGGGCTGTGCCAGCGGTGGTCACTGTCTTCTTCTTCTTCTTCTTCTTCTTCT 798
|||
Db 541 CGACACTGGGTTCTGGGCTGTGCCAGCGGTGGTCACTGTCTTCTTCTTCTTCTTCTTCTTCTTCT 600
|||
QY 799 TCTGTGATGTTTGGCTGTGAGTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCT 858
|||
Db 601 TCTGTGATGTTTGGCTGTGAGTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCT 660
|||
QY 859 CTTCTCTCTAGAACCCAGAGTGAAGTGCAGGCTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 918
|||
Db 661 CTTCTCTCTAGAACCCAGAGTGAAGTGCAGGCTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720
|||
```

```
Qy 919 GAGCAGAGCCTCGCTGAACTGTGGACCCAGACTCCGAGCCACCCAGGCGCGTGGC 978
Db 721 GAGCAGAGCCTCGCTGAACTGTGGACCCAGACTCCGAGCCACCCAGGCGCGTGGC 780
Qy 979 ACTGGTGTCAAACCTTCAACCTTCAACCTTCAACCTTCAACCTTCAACCTTGT 1038
Db 781 ACTGGTGTCAAACCTTCAACCTTCAACCTTCAACCTTCAACCTTCAACCTTGT 839
Qy 1039 TTCCATATGCGCAGACCCATAGCCGCTGCAAGGAGGAGGAGGAGGAGGAGGAGC 1098
Db 840 TTCCATATGCGCAGACCCATAGCCGCTGCAAGGAGGAGGAGGAGGAGGAGGAGC 899
Qy 1099 CTGAGTCCGACCTTGGGTGGCGGGGCTTGGGTCTCTCGTCCACCCGAGGAGCAGAC 1158
Db 900 CTGAGTCCGACCTTGGGTGGCGGGGCTTGGGTCTCTCGTCCACCCGAGGAGCAGAC 959
Qy 1159 ACCGGCTTGTGGCAGGCTGGGCTCTGTGTCAACCACTCTGTGGGTGGGTGAGACCT 1218
Db 960 ACCGGCTTGTGGCAGGCTGGGCTCTGTGTCAACCACTCTGTGGGTGGGTGAGACCT 1019
Qy 1219 TCCCTCCACCCCGGAGTCTTCAAGCTCTGCTTCTTCAAGTCTTCCAAATGGAACCC 1278
Db 1020 TCCCTCCACCCCGGAGTCTTCAAGCTCTGCTTCTTCAAGTCTTCCAAATGGAACCC 1079
Qy 1279 TCACCTCCGAGCACCAGGCTTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGC 1338
Db 1080 TCACCTCCGAGCACCAGGCTTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGC 1139
Qy 1339 CACAGACCGGACTCTCTTCTGCAACCCAGGCTGGTCCGGCCCAAGTGTGGGTGGC 1398
Db 1140 CACAGACCGGACTCTCTTCTGCAACCCAGGCTGGTCCGGCCCAAGTGTGGGTGGC 1199
Qy 1399 CTCTCTCCACTCCAGGGCTCCGCGCCCAAGTGTGGGGGCGCCCTGCGGAGCCTCAGACA 1458
Db 1200 CTCTCTCCACTCCAGGGCTCCGCGCCCAAGTGTGGGGGCGCCCTGCGGAGCCTCAGACA 1259
Qy 1459 CACTGGAGTTCAGGGC--TGGGGGGGCTTGGGACATACCTGTCCCTTGGCTATGAGCAG 1516
Db 1260 CACTGGAGTTCAGGGCTTGGGGGGGCTTGGGACATACCTGTCCCTTGGCTATGAGCAG 1319
Qy 1517 GCTTTGGGGGCTTCCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1576
Db 1320 GCTTTGGGGGCTTCCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1378
Qy 1577 CTGGGATGCTCTGCGCCCGCAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1636
Db 1379 CTGGGATGCTCTGCGCCCGCAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1437
Qy 1637 GGGGGGCTCGGCTGGGGGGTCA 1660
Db 1438 GGGGGGCTCGGCTGGGGGGTCA 1461
```

```
RESULT 9
BX384442 1060 bp mRNA linear EST 27-APR-2004
LOCUS BX384442 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
DEFINITION cDNA clone CS0DK010Y120 5-PRIME, mRNA sequence.
ACCESSION BX384442
VERSION BX384442.2 GI:46624993
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Li W.B., Gruber C., Jessee J., and Pollayes D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 8, 2003 this sequence version replaced gi:30455242.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
```

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
8574.f

For more information about this cluster, see

http://www.genoscope.cns.fr/cdna?c=CS0DK010BEL00P1&c=8574.f.

FEATURES

source

Location/Qualifiers
1..1060
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DK010Y120"
/cell_type="HELA CELLS COT 25-NORMALIZED"
/cell_line="HELA"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

	Query Match	Best Local Similarity	Score	919.6;	DB 5;	Length	1060;			
	Matches	960;	Conservative	12;	Mismatches	21;	Indels	5;	Gaps	3;
Qy	198	TTTTCTGGGGCTCGGGGCGCGGAGAGCTGCATCCACAGAGAGCGCGTCCAGAGCGGA	257							
Db	1	TTCCGGGATCTCCGGGGCGCGGAGAGCTGCATCCACAGAGAGCGCGTCCAGAGCGGA	60							
Qy	258	CCCGGAGTGTTCAGAGCCAGTGACAGAGCCAGGGGCCAAGTCCACAGCCATGC	317							
Db	61	CCCGGAGTGTTCAGAGCCAGTGACAGAGCCAGGGGCCAAGTCCACAGCCATGC	120							
Qy	318	AGACCTGCCCCCTGGCATTTCCCGCCACGTTTCCAGGCGCTTGGGACCTCTCTGT	377							
Db	121	AGACCTGCCCCCTGGCATTTCCCGCCACGTTTCCAGGCGCTTGGGACCTCTCTGT	180							
Qy	378	TGGTGTCTCTTCTGAGTGTCTCAGAAATGAAGCTGGGACAGCCCATCTTGCACAGAGGGG	437							
Db	181	TGGTGTCTCTTCTGAGTGTCTCAGAAATGAAGCTGGGACAGCCCATCTTGCACAGAGGGG	240							
Qy	438	TAGTCTCTGTCTCTTGGGGGAGAGACACCGTCATGTCTCTGCAACATCTCCAAACGCTCT	497							
Db	241	TAGTCTCTGTCTCTTGGGGGAGAGACACCGTCATGTCTCTGCAACATCTCCAAACGCTCT	300							
Qy	498	CCCATGTCAACATCAAGCTGCGTCCCAACGGGCGAGGAGCGCCATCTTCAATGAGGTGG	557							
Db	301	CCCATGTCAACATCAAGCTGCGTCCCAACGGGCGAGGAGCGCCATCTTCAATGAGGTGG	360							
Qy	558	CTCAGGCTACTTCTCCGGGACCGCTGGGAGCTCCAGGTTTCAGGGAGCGGTGGACAGC	617							
Db	361	CTCAGGCTACTTCTCCGGGACCGCTGGGAGCTCCAGGTTTCAGGGAGCGGTGGACAGC	420							
Qy	618	TGGTGTCAAAAGGCGCCCGGAGCTCCCATGTCTGGGCTGTACATGTGCACCTCTGGGAC	677							
Db	421	TGGTGTCAAAAGGCGCCCGGAGCTCCCATGTCTGGGCTGTACATGTGCACCTCTGGGAC	480							
Qy	678	ACCAGAGAAATACAGACAAAGTCAACGCTGGAGGTTTCAGGTGCAGAAACCCAGTCCGCC	737							
Db	481	ACCAGAGAAATACAGACAAAGTCAACGCTGGAGGTTTCAGGTGCAGAAACCCAGTCCGCC	540							
Qy	738	CTGACATCTGGGTTCTGGGCTGTGCAGGCTGGTCACTGTCTTCTTCTCTTGGTCTG	797							
Db	541	CCGACATCTGGGTTCTGGGCTGTGCAGGCTGGTCACTGTCTTCTTCTTCTTGGTCTG	600							
Qy	798	CTCTGGTCTATGTTCCGCTGTGTACAGGTGCGCGCTGTTCAGGCAACCGCGGAGAGAGT	857							
Db	601	CTCTGGTCTATGTTCCGCTGTGTACAGGTGCGCGCTGTTCAGGCAACCGCGGAGAGAGT	660							
Qy	858	TCCTCTCTCTAGAAACCCAGATGAAGTCTGAGCCCTTCAGAGCGGGAGCCAGCAGGGCC	917							

LOCUS BX399171 1057 bp mRNA linear EST 29-APR-2004
 DEFINITION BX399171 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
 clone CS0D1070YN15 5-PRIME, mRNA sequence.
 ACCESSION BX399171
 VERSION BX399171.2 GI:46874721
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1057)
 AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT On May 13, 2003 this sequence version replaced gi:30617803.
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
 was normalized. Library was constructed by Life technologies, a
 division of Invitrogen. This sequence belongs to sequence cluster
 8574.f
 For more information about this cluster, see
 http://www.genoscope.cns.fr/cdna?s=CS0D1070CG08QP1&c=8574.f.
 Location/Qualifiers
 1. 1057
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0D1070YN15"
 /tissue_type="PLACENTA COT 25-NORMALIZED"
 /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
 /notes="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."

FEATURES

source

ORIGIN

Query Match 40.8%; Score 889.2; DB 5; Length 1057;
 Best Local Similarity 97.4%; Pred. NO. 1.3e-199;
 Matches 956; Conservative 1; Mismatches 19; Indels 6; Gaps 5;
 QY 214 GCGCGGAGAGCTGCATCCAGAGGAGCGCTCCAGGAGCGACCCGGAGTGTTCAA 273
 DB 1 GCGCGGAGAGCTGCATCCAGAGGAGCGCTCCAGGAGCGACCCGGAGTGTTCAA 60
 QY 274 GAGCCAGTGAAGGACCGAGGCGCCAGTCCACACGACCATGCAGACCTGCCCTGGC 333
 DB 61 GAGCCAGTGAAGGACCGAGGCGCCAGTCCACACGACCATGCAGACCTGCCCTGGC 120
 QY 334 ATTCCCTGCGCCAGTTCCTCCAGGCGCTGGGACCTCTCTGTTTGGTGCCTCTTGG 393
 DB 121 ATTCCCTGCGCCAGTTCCTCCAGGCGCTGGGACCTCTCTGTTTGGTGCCTCTTGG 180
 QY 394 TGCTCAGATGAAGGCTGGGACAGCCCATCTGCACAGAGGGGTAGTCTGTGTCTTG 453
 DB 181 TGCTCAGATGAAGGCTGGGACAGCCCATCTGCACAGAGGGGTAGTCTGTGTCTTG 240
 QY 454 GGGCGAGAACCGTTCATGCTCTGCAACATCTCCACGCTTCTCCCATGTCAACATCAA 513
 DB 241 GGGCGAGAACCGTTCATGCTCTGCAACATCTCCACGCTTCTCCCATGTCAACATCAA 300
 QY 514 GCTGCTGCTCCAGCGGACGAGCGCCATCTTCAATGAGTGGCTCCAGGCTACTTCTC 573
 DB 301 GCTGCTGCTCCAGCGGACGAGCGCCATCTTCAATGAGTGGCTCCAGGCTACTTCTC 360
 QY 574 CCGGACCGCTGGCAGCTCCAGGTTTCAGGAGCGCTGGCAGACAGTGTGTCAAGAGCGC 633
 DB 361 CCGGACCGCTGGCAGCTCCAGGTTTCAGGAGCGCTGGCAGACAGTGTGTCAAGAGCGC 420

QY 634 CCGGAGCTCCCATGCTGGGCTGTATCATGTGGCACCTCGTGGACACACGAGAAATAACAG 693
 DB 421 CCGGAGCTCCCATGCTGGGCTGTATCATGTGGCACCTCGTGGACACACGAGAAATAACAG 480
 QY 694 ACAAGTCAAGCTGGAGAGTTTCAGGTGTCAGAACCCAGTCCGCCCCCTGCACACTGGGTTCG 753
 DB 481 ACAAGTCAAGCTGGAGAGTTTCAGGTGTCAGAACCCAGTCCGCCCCCTGCACACTGGGTTCG 540
 QY 754 GCTGTGCGAGGCGGTGTCATGTGCTTTCATCTCTTGTGTCGTCTGGTCAATGTCG 813
 DB 541 GCTGTGCGAGCGGTGTCATGTGCTTTCATCTCTTGTGTCGTCTGGTCAATGTCG 600
 QY 814 CTGTTACAGTGGCCGCTGTTCCCAACAGCGCGGAGAGAGTCTTCTCCCTCTAGAAC 873
 DB 601 CTGTTACAGTGGCCGCTGTTCCCAACAGCGCGGAGAGAGTCTTCTCCCTCTAGAAC 659
 QY 874 CCAGATGAAGGTGTCAGCCCTCAGAGCGGAGCCCGAGAGGCGCTTGAGCAGACCTCCGC 933
 DB 660 CCAGATGAAGGTGTCAGCCCTCAGAGCGGAGCCCGAGAGGCGCTTGAGCAGACCTCCGC 719
 QY 934 TGAAGTGTGGACCCAGACTCGAGCCCAACCCAGGCGCGTGGACTGGTGTTCAAAC 993
 DB 720 TGAAGTGTGGACCCAGACTCGAGCCCAACCCAGGCGCGTGGACTGGTGTTCAAAC 779
 QY 994 CTCACACTTGGAGCCCTGGAGCTGCTGCTCCCGCCCAACCTTGTTCATATGCGCAG 1053
 DB 780 CTCACACTTGGAGCCCTGGAGCTGCTGCTCCCGCCCAACCTTGTTCATATGCGCAG 838
 QY 1054 ACCATATGCCCTGTCAGAGGAGGAGACACAGAGAGCGCCCTGAGTGCAGACCTT 1113
 DB 839 ACCATATGCCCTGTCAGAGGAGGAGACACAGAGAGCGCCCTGAGTGCAGACCTT 897
 QY 1114 GGTGGCGGCGCTGGGTCTCTGCTCCACCGGAGGCGACAGACCGGCTTGGTGGC 1173
 DB 898 GGTGGCGGCGGCTGGGTCTCTGCTCC--ACCGGAGGCGACAGACCGGCTTGGTGGG 954
 QY 1174 AGGTGGGCGCTGCTGTGCACCC 1195
 DB 955 GGTGGGCTGTGTACCACTC 976

RESULT 12

BX358342/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Genoscope

Genoscope - Centre National de Sequencage

2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime

end enriched, double-strand cDNA was digested with Not I and cloned

into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library

was normalized. Library was constructed by Life technologies, a

division of Invitrogen. This sequence belongs to sequence cluster

8574.f

For more information about this cluster, see

http://www.genoscope.cns.fr/cdna?s=CS0D1036CD11NFI&c=8574.f.

Location/Qualifiers

1. 1140

/organism="Homo sapiens"

FEATURES

source

[illegible]

```
Db 121 TGCGTGCCTCTTGTAGTCTCAGAAATGAAGGCTGGGACAGCCCATCTGCACAGAGGGG 180
Qy 438 TAGTCTCTGTCTTGGGCGGAGAACACCGTATGTCCTTGGCAACATCTTCAAGCCTTCT 497
Db 181 TAGTCTCTGTCTTGGGCGGAGAACACCGTATGTCCTTGGCAACATCTTCAAGCCTTCT 240
Qy 498 CCATGTCATCAATCAAGCTGCGTCCGACGGGACGAGAGCGCATCTTCAATGAGGTGG 557
Db 241 CCATGTCATCAATCAAGCTGCGTCCGACGGGACGAGAGCGCATCTTCAATGAGGTGG 300
Qy 558 CTCAGGCTACTTCTCCCGGAGCGGTGGCAGCTCCAGGTTTCAGGAGCGGTGGCACAGC 617
Db 301 CTCAGGCTACTTCTCCCGGAGCGGTGGCAGCTCCAGGTTTCAGGAGCGGTGGCACAGC 360
Qy 618 TGGTGTCAAGAGCGCCCGGAGCTCCATGCTGGGCTGTACATGTGGGACCTCTGTGGAC 677
Db 361 TGGTGTCAAGAGCGCCCGGAGCTCCATGCTGGGCTGTACATGTGGGACCTCTGTGGAC 420
Qy 678 ACCAGAGAAATAACAGACAAGTCAAGCTCGAGGTTTCAGGTGCAGAACCCAGTCCGCC 737
Db 421 ACCAGAGAAATAACAGACAAGTCAAGCTCGAGGTTT-----CAGAACCCAGTCCGCC 474
Qy 738 CTGACACTGGGTTCTGGCCCTGTGCCAGCGGTGTCTACTGTCTTTCATCTTGGTGG 797
Db 475 CCGACACTGGGTTCTGGCCCTGTGCCAGCGGTGTCTACTGTCTTTCATCTTGGTGG 534
Qy 798 CTCTGTCTATGTTCCGCTGTGTACAGTGGCGGTGTCTCCAGCAACCGCGGAGAGAGT 857
Db 535 CTCTGTCTATGTTCCGCTGTGTACAGTGGCGGTGTCTCCAGCAACCGCGGAGAGAGT 594
Qy 858 TCTTCTCTCTAGAACCCAGATGAAGTGTGCAGGCTCTCAGAGCGGAGCCAGCAGGGGC 917
Db 595 TCTTCTCTCTAGAACCCAGATGAAGTGTGCAGGCTCTCAGAGCGGAGCCAGCAGGGGC 654
Qy 918 TGAGCAGAGCCTCCGCTGAACCTGTGGACCCAGACTCCGAGCCCAACCCAGCGCGTGG 977
Db 655 TGAGCAGAGCCTCCGCTGAACCTGTGGACCCAGACTCCGAGCCCAACCCAGCGCGTGG 714
Qy 978 CACTGTGTTCAAACCTCACCCTTGGAGCCCTGGAGCTGTGTCTTCCCTCCCAACCCCTTG 1037
Db 715 CACTGTGTTCAAACCTCACCCTTGGAGCCCTGGAGCTGTGT-CCCCCAACCCCTTG 773
Qy 1038 TTTCCATATGCGCAGACCAATAGCCGCTGCAAGGCAGAGAGGACACAGAGAGCCAGC 1097
Db 774 TTTCCATATGCGCAGACCAATAGCCGCTGCAA-GAGAGAGAGACACAGAGAGCCAGC 832
Qy 1098 CTGAGTGCAGACCTTTGGTGGCGGGGCTTGGTCTCTCTCTCCACCCGAGGGCAGACA 1157
Db 833 CTGAGTGCAGACCTTTGGTGGCGGGGCTTGGTCTCTCTCTCCACCCGGA-GGCACAGA 891
Qy 1158 CACCGGCTTGTGGCAGGCTGGGCTCTGTGTACCCACTCTCTGGGTGGTGGACGCC 1217
Db 892 CACAGGCTTGTGTGGCAGGCTGGG-CTYTGTTGTCA-CCACTCTGGGTGGTGGACGCT 949
Qy 1218 TTTCCCTTCCACCCCGAGTCTTCCAAAGTCTGTCTTCTTCTTCCAAATGGAAACAC 1277
Db 950 T-----CCCTTCCACCCCGAGTCTTCAAGTCTGTCTTCTTCTTCTTCAAAATGGAACT 1005
Qy 1278 CTCACCTCCGAGCAGCAGCCGAC 1298
Db 1006 YACTTCGACGACCSRCTTAC 1026
```

```
RESULT 14
BX384441/c
LOCUS
DEFINITION BX384441 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
CDNA clone CS0DK010Y120 3-PRIME, mRNA sequence.
ACCESSION BX384441
VERSION BX384441.2 GI:46620049
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
```

```
REFERENCE
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On May 8, 2003 this sequence version replaced gi:30453250.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, Cp 5706 - 91057 EVRY cedex - FRANCE
Email: segret@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
8574.f
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?8=CS0DK010BE10NP1&c=8574.f.
```

```
FEATURES
source
1..1025
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DK010Y120"
/cell_type="HELA CELLS COT 25-NORMALIZED"
/cell_line="HELA"
/clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
```

ORIGIN

```
Query Match 39.9%; Score 869.6; DB 5; Length 1025;
Best Local Similarity 94.5%; Pred. No. 5.6e-195;
Matches 976; Conservative 18; Mismatches 29; Indels 10; Gaps 9;
Qy 850 GNAGAAGTCTCTCTCTAGAACCCAGTGAAGTCTGCGCCCTCAGAGCGGGAGCCCA 909
Db 1025 GARGAATAATTTTSCBTCTARACCCAGATGAAGTTC-CAGCGCTCAGAGC-GGAGSCCA 968
Qy 910 GCAGGCGCTGACAGAGCTCCGCTGAACCTGTGGACCCAGACTCCGAGCCACCCCAAG 969
Db 967 GCAGGCGCTGACAGAG-CTCGCTGAACCTGTGMAACCCAGACTSGCAGACCCACCA 909
Qy 970 GCCGTGGCACTGTGTTCAAACTCACCCTCAGCTTGGAGCCTTGGAGCTGTCTGTCCCCC 1029
Db 908 GCGCTGGCACTGTGTTCAAACTCACCCTCAGCTTGGAGCCTGGAGCTGTCTGT-CSCCCC 850
Qy 1030 AACCTTGTTCATATGCGCAGACCCATAGCCGCTGCAAGGAGAGAGACACAGCA 1089
Db 849 AACCTTGTTCATATGCGCAGACCCATAGCCGCTGCAAGGAGAGAGACACAGCA 790
Qy 1090 GAGCAGACCCCTGAGTGGCGGACCTTGGGTGGCGGGGCTGTGGTCTCTCGTCCACCGGAG 1149
Db 789 GAGCAGACCCCTGAGTGGCGGACCTTGGGTGGCGGGGCTGTGGTCTCTCGTCCACCGGAG 730
Qy 1150 GGCACAGACACCGGCTTGTGTGGCAGGCTGGGCTCTGTGTCAACCACTCTGTGGGTGCT 1209
Db 729 GGCACAGACACCGGCTTGTGTGGCAGGCTGGGCTCTGTGTCAACCACTCTGTGGGTGCT 670
Qy 1210 GCAGACCCCTTCCCTTCCACCCCGAGTCTTCCAGTCTGTCTCTCAGTTTCCAAAT 1269
Db 669 GCAGACCCCTTCCCTTCCACCCCGAGTCTTCCAGTCTGTCTCTCAGTTTCCAAAT 610
Qy 1270 GGAAACACCTCACTCCGAGCAGCCGACTTACAGGAGCGCATGCCCTTCCCTCTGCCCT 1329
Db 609 GGAAACACCTCACTCCGAGCAGCCGACTTACAGGAGCGCATGCCCTTCCCTCTGCCCT 550
Qy 1330 CATCAAAACCCAGACCCGGAATCCCTTTCTGCCACCCAGGCTGTCTCGGCGCCAGGTG 1389
Db 549 CATCAAAACCCAGACCCGGAATCCCTTTCTGCCACCCAGGCTGTCTCGGCGCCAGGTG 490
```


Db 896 ACCTGGGTGGGGG-CTGGGTCTCTGTCACCGGAGGCACAGACACGGCTTGCTTG 954
Qy 1169 TTGGCAGGCT 1178
Db 955 CAGGTGGGCT 964

Search completed: March 20, 2005, 06:49:48
Job time : 7213 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 20, 2005, 09:49:49 ; Search time 85 Seconds
(without alignments)
1128.430 Million cell updates/sec

Title: US-10-080-522-1

Perfect score: 248

Sequence: 1 MQTCPLAFPGHVSQLGTL.....PLGALELLSQPLFPFYAADP 248

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 2105692 seqs, 386760381 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

- 1: geneseqp1980s:*
- 2: geneseqp1990s:*
- 3: geneseqp2000s:*
- 4: geneseqp2001s:*
- 5: geneseqp2002s:*
- 6: geneseqp2003as:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	248	100.0	248	2 AAW48811	AAW48811 K12 prote
2	248	100.0	248	2 AAY21846	AAY21846 Human sig
3	248	100.0	248	4 AAB36658	AAB36658 Human K12
4	248	100.0	248	8 ADJ75366	ADJ75366 Marker ge
5	248	100.0	248	8 ADP24664	ADP24664 PRO poly
6	136	54.8	149	7 ADB36336	ADB36336 Human imm
7	134	54.0	162	6 AAO29896	AAO29896 Human org
8	123	49.6	183	6 AAO29897	AAO29897 Human org
9	92	37.1	101	6 AAO29895	AAO29895 Human org
10	55	22.2	55	3 AAB34671	AAB34671 Gene 35 h
11	55	22.2	107	6 AAO29898	AAO29898 Human org
12	48	19.4	55	3 AAB34672	AAB34672 Human sec
13	31	12.5	43	7 ADB36337	ADB36337 Human imm
14	30	12.1	30	3 AAB34673	AAB34673 Gene 35 h
15	12	4.8	27	3 AAB34674	AAB34674 Human sec
16	8	3.2	294	6 ABU20614	ABU20614 Protein e
17	8	3.2	340	8 ABM80522	ABM80522 Tumour-as
18	8	3.2	384	6 ABU48570	ABU48570 Protein e
19	8	3.2	453	7 ABO82888	ABO82888 Pseudomon
20	8	3.2	691	8 ADE21140	ADE21140 Bacterial
21	7	2.8	9	8 ADK03559	ADK03559 Hepatitis
22	7	2.8	10	8 ADK03571	ADK03571 Hepatitis
23	7	2.8	17	5 AAE21241	AAE21241 Human gen
24	7	2.8	36	4 AAE01367	AAE01367 Human gen
25	7	2.8	36	5 ABG64111	ABG64111 Human alb

26 7 2.8 36 8 ADL77376
27 7 2.8 40 3 AAG57017
28 7 2.8 56 8 ABO55459
29 7 2.8 58 6 ABP77249
30 7 2.8 61 8 ADK01568
31 7 2.8 67 5 ABP10379
32 7 2.8 67 5 ABP32516
33 7 2.8 69 3 AAB45118
34 7 2.8 70 4 AAU43324
35 7 2.8 70 6 ABM39843
36 7 2.8 72 6 ADA21112
37 7 2.8 91 2 AAY48485
38 7 2.8 108 4 AAG73477
39 7 2.8 110 3 AAB40620
40 7 2.8 118 3 AAB45071
41 7 2.8 123 4 AAU48594
42 7 2.8 123 6 ABM45113
43 7 2.8 127 8 ADJ67024
44 7 2.8 128 7 ADB64958
45 7 2.8 128 7 ADM03909

ALIGNMENTS

RESULT 1
AAW48811
ID AAW48811 standard; protein; 248 AA.
XX
AC AAW48811;
XX
DT 26-OCT-1998 (first entry)
XX
DE K12 protein.
XX
KW DNA probe; CD7 HSI DNase hypersensitive site; mRNA northern blot;
KW human erythroleukemic; HEL; K562 cell line; clone; breast cancer;
KW ovarian cancer; malignant; tumour.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 239
FT FT /note= "encoded by CCAA"
XX
PN WO9822502-A1.
XX
PD 28-MAY-1998.
XX
PF 24-NOV-1997; 97WO-US021517.
XX
PR 22-NOV-1996; 96US-00755559.
XX
PA (UYDU-) UNIV DUKE.
XX
PI Kaufman RE, Slentz-Kesler KA;
XX
DR WPI; 1998-312415/27.
XX
DR N-PSDB; AAV32446.
XX
PT New isolated K12 protein gene - which is over expressed in certain
PT neoplastic cells, used to develop products for tumour detection and
PT treatment.
XX
PS Claim 3; Fig 1; 44pp; English.
XX
CC This present sequence represents the K12 protein, the gene for which has
CC been located on chromosome 17q25. To obtain this genes cDNA sequence a
CC 500 bp DNA probe, which can be located just upstream of the CD7 HSI DNase
CC hypersensitive site, was used against a mRNA northern blot. From this a
CC 1.8kb transcript was detected in the human erythroleukemic cell line HEL.
CC The probe was then used to screen a human erythroleukemic cell line K562
CC cDNA library, from which several clones were identified and isolated that

Human signal-peptide containing protein coding sequences used to treat cancer and immune responses.

Claim 1; Page 79-80; 99pp; English.

The invention provides human signal-peptide containing proteins (SIGP) (AAY21841-855) and polynucleotides (AAK82076-90) encoding the proteins. A host cell containing a vector comprising SIGP DNA can be used to produce the SIGP protein. The SIGP protein can be used, in conjunction with a pharmaceutical carrier to treat or prevent a cancer. An antagonist of the SIGP protein can be used to treat or prevent a cancer or an immune response. The cancers that can be treated or prevented include sarcomas, adenocarcinomas, leukemias, lymphomas, melanomas, teratocarcinomas, myelomas and cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus. The immune responses that can be treated or prevented include, AIDS, Addison's disease, adult respiratory distress syndrome, allergies, anemia, asthma, atherosclerosis, bronchitis, cholecystitis, Crohn's disease, ulcerative colitis, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, atrophic gastritis, glomerulonephritis, Grave's disease, gout, hyperosinophilia, irritable bowel syndrome, lupus erythematosus, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, rheumatoid arthritis, scleroderma, Sjogren's syndrome, and autoimmune thyroiditis, complications of cancer, infections, and trauma

Sequence 248 AA;

Query Match 100.0%; Score 248; DB 2; Length 248;
Best Local Similarity 100.0%; Pred. No. 2.5e-242;
Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MOTCPAPPHGVHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSGENTVMSNCISNA 60
DB 1 MOTCPAPPHGVHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSGENTVMSNCISNA 60

QY 61 FSHVNIKLRAHQESAIFFNEVAPGYFSRDGWLQVGGVAQLVTKGARDSHAGLYMHLV 120
DB 61 FSHVNIKLRAHQESAIFFNEVAPGYFSRDGWLQVGGVAQLVTKGARDSHAGLYMHLV 120

QY 121 GHORNNRQVTLVSGAEPQSDPTGFWPVPVAVTVFILLVALVNFAYRCRCQORREK 180
DB 121 GHORNNRQVTLVSGAEPQSDPTGFWPVPVAVTVFILLVALVNFAYRCRCQORREK 180

QY 181 KFFLEPQMKVAALRAGAQQGLSRASAEIWTDPSEPTPRPLALVFKPSPLGALLELLSPQ 240
DB 181 KFFLEPQMKVAALRAGAQQGLSRASAEIWTDPSEPTPRPLALVFKPSPLGALLELLSPQ 240

QY 241 LFPYAADP 248
DB 241 LFPYAADP 248

RESULT 3
AAB36658
ID AAB36658 standard; protein; 248 AA.
XX AAB36658;
AC AAB36658;
XX 13-MAR-2001 (first entry)
DT 13-MAR-2001 (first entry)
DE Human K12 protein sequence SEQ ID NO:4.
XX Human; CD7; K12; cognate ligand; cluster of differentiation; cancer;
KW identification; inhibiting T cell proliferation; HIV; infection;
KW activating natural killer cell proliferation; leukaemia; lymphoma;
KW sepsis; graft versus host disease; autoimmune disease; arthritis;
KW multiple sclerosis; rheumatoid arthritis; psoriatic arthritis; lupus;
KW scleroderma; psoriasis; atopic dermatitis; type 1 diabetes mellitus;
KW Hashimoto's thyroiditis; pernicious anaemia; Addison's disease; uveitis;
KW myasthenia gravis; psoriasis; Guillain-Barre Syndrome; Grave's disease;

constituted a 1.8kb cDNA. This cDNA was designated K12 and was found to have a single open reading frame as well as being in the same orientation as CD7. The K12 gene was found to be expressed in both breast and ovarian cancer cells at a much higher level than any other malignant or normal tissue that was examined, thus enabling the K12 to be a useful protein in tumour detection and treatment

Sequence 248 AA;

Query Match 100.0%; Score 248; DB 2; Length 248;
Best Local Similarity 100.0%; Pred. No. 2.5e-242;
Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MOTCPAPPHGVHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSGENTVMSNCISNA 60
DB 1 MOTCPAPPHGVHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSGENTVMSNCISNA 60

QY 61 FSHVNIKLRAHQESAIFFNEVAPGYFSRDGWLQVGGVAQLVTKGARDSHAGLYMHLV 120
DB 61 FSHVNIKLRAHQESAIFFNEVAPGYFSRDGWLQVGGVAQLVTKGARDSHAGLYMHLV 120

QY 121 GHORNNRQVTLVSGAEPQSDPTGFWPVPVAVTVFILLVALVNFAYRCRCQORREK 180
DB 121 GHORNNRQVTLVSGAEPQSDPTGFWPVPVAVTVFILLVALVNFAYRCRCQORREK 180

QY 181 KFFLEPQMKVAALRAGAQQGLSRASAEIWTDPSEPTPRPLALVFKPSPLGALLELLSPQ 240
DB 181 KFFLEPQMKVAALRAGAQQGLSRASAEIWTDPSEPTPRPLALVFKPSPLGALLELLSPQ 240

QY 241 LFPYAADP 248
DB 241 LFPYAADP 248

RESULT 2
AAY21846
ID AAY21846 standard; protein; 248 AA.
XX AAY21846;
XX 20-SEP-1999 (first entry)
DE Human signal peptide-containing protein (SIGP) (clone ID 1747327).
XX Signal-peptide containing protein; SIGP; human; cancer; immune response;
KW adenocarcinoma; leukemia; lymphoma; melanoma; myeloma; sarcoma; AIDS;
KW Addison's disease; adult respiratory distress syndrome; allergy; anemia;
KW asthma; atherosclerosis; bronchitis; cholecystitis; Crohn's disease;
KW ulcerative colitis; atopic dermatitis; dermatomyositis; emphysema;
KW diabetes mellitus; atrophic gastritis; glomerulonephritis; gout; trauma;
KW Grave's disease; hyperosinophilia; irritable bowel syndrome; infection;
KW lupus erythematosus; multiple sclerosis; myasthenia gravis; inflammation;
KW osteoarthritis; osteoporosis; pancreatitis; polymyositis; scleroderma;
KW rheumatoid arthritis; Sjogren's syndrome; autoimmune thyroiditis.
XX Homo sapiens.
OS
XX W09933981-A2.
XX 08-JUL-1999.
XX 22-DEC-1998; 98WO-US027598.
XX 31-DEC-1997; 97US-00002485.
XX (INCY-) INCYTE PHARM INC.
XX Lal P, Hillman JL, Corley NC, Guegler KJ, Baughn MR, Sather SK;
XX Shah P;
XX WPI; 1999-430242/36.
XX N-PSDB; AAX82081.
XX

KW systemic lupus erythematosus; dermatomyositis; asthma; eczema;
KW atopic dermatitis; contact dermatitis; eczematous dermatitis;
KW seborrheic dermatitis; rhinitis.

OS Homo sapiens.

XX WO200073333-A2.

XX 07-DEC-2000.

XX 26-MAY-2000; 2000WO-US014612.

XX 28-MAY-1999; 99US-0136450P.

XX (IMV) IMMUNEX CORP.

XX Lyman SD, Fanslow WC;

XX WPI; 2001-061511/07.

DR N-PSDB; AAC88152.

XX Stimulating intracellular signaling of CD7 comprises contacting a cell
PT expressing CD7 with recombinant K12 protein, the cognate ligand of CD7,
PT to inhibit T cell proliferation and/or activate natural killer cell
PT proliferation.

XX Claim 2; Page 38-39; 42pp; English.

XX The present invention describes a method for stimulating (S) the
CC intracellular signalling of CD (cluster of differentiation) 7 comprising
CC contacting a cell that expresses CD7 with a recombinant K12 protein (I),
CC the cognate ligand of CD7. (S) is useful for inhibiting T cell
CC proliferation and/or activating NK (natural killer) cell proliferation
CC and/or inducing NK toxicity in a mammal which involves administration of
CC K12 protein. It is also used for treating HIV-1 infection, cancer (T cell
CC leukaemia, acute lymphocytic leukaemia, cutaneous T cell lymphoma),
CC bacterial and viral infections, mediated by CD7. In the case of treating
CC T cell leukaemia the soluble K12 protein is covalently attached to a
CC toxin. A disease mediated by CD7 such as sepsis, graft versus host
CC disease due to transplantation, autoimmune diseases, multiple sclerosis,
CC arthritis, rheumatoid arthritis, psoriatic arthritis, scleroderma, lupus,
CC psoriasis, atopic dermatitis, type 1 diabetes mellitus, Hashimoto's
CC thyroiditis, pernicious anaemia, Addison's disease, myasthenia gravis,
CC uveitis, psoriasis, Guillain-Barre Syndrome, Grave's disease, systemic
CC lupus erythematosus and dermatomyositis, asthma, eczema, atopic
CC dermatitis, contact dermatitis, other eczematous dermatides, seborrheic
CC dermatitis, and rhinitis is also treated by administering a K12
CC antagonist (neutralising antibody). The present sequence represents the
CC human K12 protein, which is given in the exemplification of the present
CC invention

XX Sequence 248 AA;

Query Match 100.0%; Score 248; DB 4; Length 248;
Best Local Similarity 100.0%; Pred. No. 2.5e-242;
Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTCPLAPGHVSQALGTLFLAASLSAQNEGWDSPICTEGVVSVSGENTVMSNCISNA 60
DB 1 MTCPLAPGHVSQALGTLFLAASLSAQNEGWDSPICTEGVVSVSGENTVMSNCISNA 60
QY 61 FSHVNIKLRAHQESAIENEVAPGYFSRDGMQLQVGGVAQLVIKARDSHAGLYMHLV 120
DB 61 FSHVNIKLRAHQESAIENEVAPGYFSRDGMQLQVGGVAQLVIKARDSHAGLYMHLV 120
QY 121 GHQRNRRQVTLVSGAEPQSDPTGFWPVAVTAVFILLVALVNFAYWRCRCQORREK 180
DB 121 GHQRNRRQVTLVSGAEPQSDPTGFWPVAVTAVFILLVALVNFAYWRCRCQORREK 180
QY 181 KFFLLEPQMKVAALRAGAQGGLSASAEIWTDPSEPTPRPLALVFKPSPGLALELLSPQ 240
DB 181 KFFLLEPQMKVAALRAGAQGGLSASAEIWTDPSEPTPRPLALVFKPSPGLALELLSPQ 240

QY

241 LFPYAADP 248

|||||||

Db 241 LFPYAADP 248

RESULT 4

ADJ75366

ID ADJ75366 standard; protein; 248 AA.

XX AC ADJ75366;

XX 20-MAY-2004 (first entry)

XX Marker gene related amino acid sequence SEQ ID NO:618.

XX bronchial asthma; chronic obstructive pulmonary disease;

XX respiratory epithelial cell; interleukin-13; respiratory; antiasthmatic;
KW gene therapy; marker.

XX OS Homo sapiens.

XX EPI394274-A2.

XX 03-MAR-2004.

XX 04-AUG-2003; 2003EP-00254857.

XX 06-AUG-2003; 2002JP-00229312.

XX 20-MAR-2003; 2003JP-00077212.

XX (GENO-) GENOX RES INC.

XX Ohtani N, Sugita Y, Yamaya M, Kubo H, Nagai H, Izuhara K;

XX WPI; 2004-193155/19.

XX Testing for bronchial asthma or chronic obstructive pulmonary disease by
PT comparing the expression level of a marker gene in a biological sample
PT from a subject with the expression level of the gene in a sample from a
PT healthy subject.

XX Example 11; SEQ ID NO 618; 241pp; English.

XX The present invention describes a method of testing for bronchial asthma
CC or chronic obstructive pulmonary disease. The method comprises
CC determining the expression level of a marker gene in a biological sample
CC from a subject, comparing the expression level determined with the
CC expression level of the marker gene in a biological sample from a healthy
CC subject, and judging whether the subject has bronchial asthma or chronic
CC obstructive pulmonary disease. The marker gene comprises: (a) a group of
CC genes (S1) whose expression levels increase when respiratory epithelial
CC cells are stimulated with interleukin-13; or (b) a group of genes (S2)
CC whose expression levels decrease when respiratory epithelial cells are
CC stimulated with interleukin-13. Also described: (1) a reagent (I) for
CC testing for bronchial asthma or chronic obstructive pulmonary disease;
CC (2) a kit for screening for a candidate compound for a therapeutic agent
CC to treat bronchial asthma or chronic obstructive pulmonary disease; (3)
CC an animal model for bronchial asthma or chronic obstructive pulmonary
CC disease; (4) an inducer that induces bronchial asthma in a mouse; (5) a
CC method for producing an animal model for bronchial asthma or chronic
CC obstructive pulmonary disease; (6) a therapeutic agent for bronchial
CC asthma or chronic obstructive pulmonary disease, comprising the compound,
CC a marker gene or an antisense nucleic acid corresponding to a portion of
CC the marker gene, a ribozyme, a polynucleotide that suppresses the
CC expression of the gene through an RNAi effect or an antibody recognising
CC a protein encoded by a marker gene; and (7) a DNA chip for testing for
CC bronchial asthma or a chronic obstructive pulmonary disease, on which a
CC probe has been immobilised to assay a marker gene. (I) has respiratory
CC and antiasthmatic activities, and can be used in gene therapy. The method
CC is useful for testing for or screening for a therapeutic agent for
CC bronchial asthma or chronic obstructive pulmonary disease. The present
CC sequence is used in the exemplification of the present invention.

XX

SQ Sequence 248 AA;
 Query Match 100.0%; Score 248; DB 8; Length 248;
 Best Local Similarity 100.0%; Pred. No. 2.5e-242; Mismatches 0; Indels 0; Gaps 0;
 Matches 248; Conservative 0;
 QY 1 MQTCLAPFGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSNCISNA 60
 DB 1 MQTCLAPFGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSNCISNA 60
 QY 61 FSHVNIKLRAHQESAI FNEVAPGYSRDGQVQVGVQVQLVKGARDSHAGLYMHWLV 120
 DB 61 FSHVNIKLRAHQESAI FNEVAPGYSRDGQVQVGVQVQLVKGARDSHAGLYMHWLV 120
 QY 121 GHORNRQVTLVSGAEPOSAPDTGFWPVAVTAVFILLVAVMFAYWRCRCQORREK 180
 DB 121 GHORNRQVTLVSGAEPOSAPDTGFWPVAVTAVFILLVAVMFAYWRCRCQORREK 180
 QY 181 KFFLLEPQMKVAALRAGAQQGLSRASAEI WTPDSEPTPRPLALVFKPSPLGALLELLSPQ 240
 DB 181 KFFLLEPQMKVAALRAGAQQGLSRASAEI WTPDSEPTPRPLALVFKPSPLGALLELLSPQ 240
 QY 241 LFPYAADP 248
 DB 241 LFPYAADP 248
 RESULT 5
 ADP24664
 ID ADP24664 standard; protein; 248 AA.
 XX
 AC ADP24664;
 DT 18-NOV-2004 (first entry)
 XX
 DE PRO polypeptide SEQ ID NO:1842.
 XX
 KW PRO; antiinflammatory; antiarthritic; antirheumatic; immunosuppressive;
 KW osteopathic; antidiabetic; dermatological; antipsoriatic; antiallergic;
 KW antiasthmatic; hepatotropic; respiratory; gene therapy; immune system.
 XX
 OS Unidentified.
 XX
 PN WO2004041170-A2.
 XX
 PD 21-MAY-2004.
 XX
 PF 30-OCT-2003; 2003WO-US034312.
 XX
 PR 01-NOV-2002; 2002US-0423394P.
 XX
 PA (GETH) GENENTECH INC.
 XX
 PI Clark H, Schoenfeld J, Van Lookeren M, Williams PM, Wood WI;
 PI Wu TD;
 XX
 DR WPI; 2004-419628/39.
 DR N-PSDB; ADP24663.
 XX
 XX New PRO polypeptides and polynucleotides, useful for treating e.g.
 PT erythematous, rheumatoid arthritis, diabetes mellitus, immune-mediated
 PT renal disease, or demyelinating diseases of the central or peripheral
 PT nervous system.
 XX
 PS Claim 7; SEQ ID NO 1842; 2940pp; English.
 XX
 CC The invention relates to a novel isolated nucleic acid and the PRO
 CC polypeptide encoded by it. A protein of the invention has
 CC antiinflammatory, antiarthritic, antirheumatic, immunosuppressive,
 CC osteopathic, antidiabetic, dermatological, antipsoriatic, antiallergic,
 CC antiasthmatic, hepatotropic, and respiratory activity. A polynucleotide
 CC of the invention may have a use in gene therapy. The PRO polypeptide, its
 CC agonist, antagonist, or antibody that specifically binds to the

CC polypeptide is useful for treating an immune related disorder such as
 CC systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis,
 CC juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis, an
 CC idiopathic inflammatory myopathy, Sjogren's syndrome, autoimmune
 CC vasculitis, sarcoidosis, autoimmune haemolytic anaemia, systemic
 CC thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal
 CC disease, a demyelinating disease of the central or peripheral nervous
 CC system, idiopathic demyelinating polyneuropathy, Guillain-Barre syndrome,
 CC a chronic inflammatory demyelinating polynuropathy, a hepatobiliary
 CC disease, infectious or autoimmune chronic active hepatitis, primary
 CC biliary cirrhosis, granulomatous hepatitis, sclerosing cholangitis,
 CC inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's
 CC disease, an autoimmune or immune-mediated skin disease, a bullous skin
 CC disease, erythema multiforme, contact dermatitis, psoriasis, an allergic
 CC disease, asthma, allergic rhinitis, atopic dermatitis, food
 CC hypersensitivity, urticaria, an immunologic disease of the lung,
 CC eosinophilic pneumonia, idiopathic pulmonary fibrosis, hypersensitivity
 CC pneumonitis, a transplantation associated disease, graft rejection or
 CC graft-versus-host disease. The present sequence represents a PRO protein
 CC of the invention.
 XX
 SQ Sequence 248 AA;
 Query Match 100.0%; Score 248; DB 8; Length 248;
 Best Local Similarity 100.0%; Pred. No. 2.5e-242; Mismatches 0; Indels 0; Gaps 0;
 Matches 248; Conservative 0;
 QY 1 MQTCLAPFGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSNCISNA 60
 DB 1 MQTCLAPFGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSNCISNA 60
 QY 61 FSHVNIKLRAHQESAI FNEVAPGYSRDGQVQVGVQVQLVKGARDSHAGLYMHWLV 120
 DB 61 FSHVNIKLRAHQESAI FNEVAPGYSRDGQVQVGVQVQLVKGARDSHAGLYMHWLV 120
 QY 121 GHORNRQVTLVSGAEPOSAPDTGFWPVAVTAVFILLVAVMFAYWRCRCQORREK 180
 DB 121 GHORNRQVTLVSGAEPOSAPDTGFWPVAVTAVFILLVAVMFAYWRCRCQORREK 180
 QY 181 KFFLLEPQMKVAALRAGAQQGLSRASAEI WTPDSEPTPRPLALVFKPSPLGALLELLSPQ 240
 DB 181 KFFLLEPQMKVAALRAGAQQGLSRASAEI WTPDSEPTPRPLALVFKPSPLGALLELLSPQ 240
 QY 241 LFPYAADP 248
 DB 241 LFPYAADP 248
 RESULT 6
 ADB36336
 ID ADB36336 standard; protein; 149 AA.
 XX
 AC ADB36336;
 DT 04-DEC-2003 (first entry)
 XX
 DE Human immune response associated protein IRAP-7 SEQ ID NO:7.
 XX
 KW human; immune response associated protein; IRAP; anti-HIV; antiallergic;
 KW antianemic; antiasthmatic; antiarteriosclerotic; dermatological;
 KW antiinflammatory; antidiabetic; nephrotropic; antihypertrophic; thyromimetic;
 KW immunosuppressive; antirheumatic; antiarthritic; osteopathic;
 KW antibacterial; virucide; antiparasitic; protozoacide; fungicide;
 KW cerebroprotective; neuroprotective; nootropic; antiparkinsonian;
 KW antipsoriatic; cytostatic; cardiant; gene therapy;
 KW immune system disorder; neurological disorder; developmental disorder;
 KW muscle disorder; cell proliferative disorder; AIDS; allergy; anaemia;
 KW asthma; atherosclerosis; contact dermatitis; Crohn's disease; diabetes;
 KW glomerulonephritis; Grave's disease; Hashimoto's thyroiditis;
 KW multiple sclerosis; rheumatoid arthritis; osteoporosis;
 KW systemic lupus erythematosus; infection; stroke; Alzheimer's disease;
 KW Parkinson's disease; psoriasis; cancer; cardiomyopathy.
 XX

OS Homo sapiens.
 XX WO2003074726-A2.
 XX 12-SEP-2003.
 XX 28-FEB-2003; 2003WO-US006307.
 XX 01-MAR-2002; 2002US-0361088P.
 XX 27-MAR-2002; 2002US-0368494P.
 XX 10-MAY-2002; 2002US-0379876P.
 XX 28-JUN-2002; 2002US-0392641P.
 XX (INCY-) INCYTE GENOMICS INC.
 XX Chawla NK, Becha SD, Jin P, Tran UK, Kable AE, Lee S;
 PI Marquis JP, Emerling BM, Ison CH, Ramkumar J, Chang H;
 XX WPI; 2003-722079/68.
 DR N-PSDB; ADB36352.
 XX New human immune response associated proteins and polynucleotides, useful
 PT for diagnosing, preventing or treating diseases or conditions associated
 PT with aberrant enzyme expression, e.g. cancer, AIDS, atherosclerosis or
 PT stroke.
 XX Claim 1; Page 143; 158pp; English.
 XX ADB36346 to ADB36361 encode the human immune response associated proteins
 CC given in ADB36330 to ADB36345, designated IRAP-1 to IRAP-16. IRAP
 CC sequences have anti-HIV, antiallergic, antianemic, antiasthmatic,
 CC antiarteriosclerotic, dermatological, antiinflammatory, antidiabetic,
 CC nephrotoxic, antithyroid, thyromimetic, immunosuppressive,
 CC antirheumatic, antiarthritic, osteopathic, antibacterial, virucide,
 CC antiparasitic, protozoacide, fungicide, cerebroprotective,
 CC neuroprotective, nootropic, antiparkinsonian, antipsoriatic, cytostatic
 CC and cardiant activities, and can be used in gene therapy. The IRAP
 CC proteins and polynucleotides can be used in diagnosing, preventing or
 CC treating diseases or conditions associated with the decreased expression
 CC or overexpression of IRAP, such as immune system, neurological,
 CC developmental, muscle or cell proliferative disorders. The disorders may
 CC include AIDS, allergies, anaemia, asthma, atherosclerosis, contact
 CC dermatitis, Crohn's disease, diabetes, glomerulonephritis, Grave's
 CC disease, Hashimoto's thyroiditis, multiple sclerosis, rheumatoid
 CC arthritis, osteoporosis, systemic lupus erythematosus, infections (e.g.
 CC bacterial, viral, fungal, parasitic or protozoal), stroke, Alzheimer's
 CC disease, Parkinson's disease, psoriasis, cancer and cardiomyopathy. They
 CC are also useful in assessing the effects of exogenous compounds on the
 CC expression of nucleic acid and amino acid sequences of IRAP. The IRAP or
 CC its fragments are useful in screening compounds for effectiveness as
 CC agonist or antagonist of the polypeptides, or in altering the expression
 CC of the target polynucleotide and compounds that specifically bind to or
 CC modulate the activity of the polypeptide.
 XX Sequence 149 AA;
 SQ
 Query Match 54.8%; Score 136; DB 7; Length 149;
 Best Local Similarity 100.0%; Pred. No. 4.2e-129;
 Matches 136; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MOTCPAPFGHVSQALGTLLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSNCISNA 60
 Db 1 MOTCPAPFGHVSQALGTLLFLAASLSAQNEGWDSPICTEGVSVSVSWGENTVMSNCISNA 60
 QY 61 FSHVNIKLRHGOEAIENEVAPGYFRDQWLOVQGGVAQLVKGARDSHAGLYWHV 120
 Db 61 FSHVNIKLRHGOEAIENEVAPGYFRDQWLOVQGGVAQLVKGARDSHAGLYWHV 120
 QY 121 GHQRNNRQVTLVSGA 136
 Db 121 GHQRNNRQVTLVSGA 136

RESULT 7
 AAO29896
 ID AAO29896 standard; protein; 162 AA.
 XX
 AC AAO29896;
 XX
 DT 03-SEP-2003 (first entry)
 XX
 DE Human organelle-associated protein (ORGA)-9.
 XX
 KW Human; organelle-associated protein; ORGA; cell proliferative disorder;
 KW actinic keratosis; arteriosclerosis; reproductive disorder; infertility;
 KW endometriosis; gastrointestinal disorder; ectopic pregnancy; hepatitis;
 KW gastroenteritis; neurological disorder; Alzheimer's disease; cirrhosis;
 KW glomerulonephritis; renal amyloidosis; renal failure; Addison's disease;
 KW renal disorder; ovulatory defect; teratogenesis; Pick's disease; cancer;
 KW Huntington's disease; urologic disorder; cystitis; dementia; dysphagia;
 KW indigestion; gastritis; hypertension; gene therapy; bursitis; anorexia;
 KW infection; urethritis; prostatitis.
 XX
 OS Homo sapiens.
 XX
 FN WO2003044171-A2.
 XX
 PD 30-MAY-2003.
 XX
 PF 15-NOV-2002; 2002WO-US036807.
 XX
 PR 16-NOV-2001; 2001US-0332384P.
 PR 13-DEC-2001; 2001US-0341187P.
 PR 23-JAN-2002; 2002US-0351151P.
 PR 27-FEB-2002; 2002US-0360269P.
 PR 05-APR-2002; 2002US-0370637P.
 PR 14-JUN-2002; 2002US-0388946P.
 XX (INCY-) INCYTE GENOMICS INC.
 PA
 PI Ramkumar J, Burford N, Xu Y, Chawla NK, Kable AE, Becha SD;
 PI Richardson TW, Gorvad AE, Sprague WW, Elliott VS, Khare R;
 PI Hawkins PR, Jin P, Lee SY, Marquis JP, Swarnakar A, Wilson AD;
 PI Hafalia AJA, Bulloch S;
 XX WPI; 2003-457603/43.
 DR N-PSDB; AAL60549.
 XX
 PT New organelle-associated proteins and polynucleotides, useful for
 PT diagnosing, treating and/or preventing cell proliferative, reproductive,
 PT gastrointestinal, neurological, urologic, and renal disorders.
 XX
 PS Claim 1; Page 166; 194pp; English.
 XX
 CC The invention relates to human organelle-associated proteins (ORGA) and
 CC nucleic acid molecules encoding such proteins. ORGA sequences are useful
 CC for diagnosing, treating and/or preventing cell proliferative disorders
 CC (e.g. actinic keratosis, arteriosclerosis, bursitis, cirrhosis, cancers
 CC or hepatitis), reproductive disorders (e.g. infertility, endometriosis,
 CC ovulatory defects, ectopic pregnancy or teratogenesis), gastrointestinal
 CC disorders (e.g. dysphagia, peptide oesophagitis, gastritis, indigestion,
 CC anorexia or gastroenteritis), neurological disorders (e.g. Alzheimer's
 CC disease, Pick's disease, Huntington's disease or dementia), urologic
 CC disorders (e.g. infection, urethritis, cystitis, prostatitis or upper
 CC tract infections) and renal disorders (e.g. glomerulonephritis, renal
 CC amyloidosis, renal failure, Addison's disease or hypertension). The
 CC invention is also useful in gene therapy. The present sequence is human
 CC ORGA protein
 XX
 SQ Sequence 162 AA;
 Query Match 54.0%; Score 134; DB 6; Length 162;
 Best Local Similarity 100.0%; Pred. No. 4.8e-127;
 Matches 134; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MOTCPAPFGHVSQALGTLLFLAASLSAQNEGWDSPICTEGVSVSVSWGENTVMSNCISNA 60

CC tract infections) and renal disorders (e.g. glomerulonephritis, renal
 CC amyloidosis, renal failure, Addison's disease or hypertension). The
 CC invention is also useful in gene therapy. The present sequence is human
 CC ORGA protein
 XX
 SQ Sequence 183 AA;
 Query Match 49.6%; Score 123; DB 6; Length 183;
 Best Local Similarity 100.0%; Pred. No. 7.4e-116;
 Matches 123; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MQTCPLAFPGHVSQALGTLFLAASLSAQNEGWDSPICTEGVVSVSGENTVMSCNISNA 60
 DB 1 MQTCPLAFPGHVSQALGTLFLAASLSAQNEGWDSPICTEGVVSVSGENTVMSCNISNA 60
 QY 61 FSHVNIKLRAHQESAIENEVAPGYFSDGWLQVGGVAQLVIKGARDSHAGLYMHLV 120
 DB 61 FSHVNIKLRAHQESAIENEVAPGYFSDGWLQVGGVAQLVIKGARDSHAGLYMHLV 120
 QY 121 GHQ 123
 DB 121 GHQ 123
 RESULT 9
 ID AAO29895 standard; protein; 101 AA.
 XX AAO29895;
 AC AAO29895;
 DT 03-SEP-2003 (first entry)
 XX Human organelle-associated protein (ORGA)-8.
 DE Human; organelle-associated protein; ORGA; cell proliferative disorder;
 XX actinic keratosis; arteriosclerosis; reproductive disorder; infertility;
 KW endometriosis; gastrointestinal disorder; ectopic pregnancy; hepatitis;
 KW gastroenteritis; neurological disorder; Alzheimer's disease; cirrhosis;
 KW glomerulonephritis; renal amyloidosis; renal failure; Addison's disease;
 KW renal disorder; ovulatory defect; teratogenesis; Pick's disease; cancer;
 KW Huntington's disease; urologic disorder; cystitis; dementia; dysphagia;
 KW indigestion; gastritis; hypertension; gene therapy; bursitis; anorexia;
 KW infection; urethritis; prostatitis.
 XX Homo sapiens.
 OS WO2003044171-A2.
 PN 30-MAY-2003.
 PD 15-NOV-2002; 2002WO-US036807.
 XX 16-NOV-2001; 2001US-0332384P.
 PR 13-DEC-2001; 2001US-0341187P.
 PR 23-JAN-2002; 2002US-0351151P.
 PR 27-FEB-2002; 2002US-0360269P.
 PR 05-APR-2002; 2002US-0370637P.
 PR 14-JUN-2002; 2002US-0388946P.
 XX (INCY-) INCYTE GENOMICS INC.
 PA Rankumar J, Burford N, Xu Y, Chawla NK, Kable AE, Becha SD;
 PI Richardson TW, Gorvad AE, Sprague WW, Elliott VS, Khare R;
 PI Hawkins PR, Jin P, Lee SY, Marquis JP, Swarnakar A, Wilson AD;
 PI Hafalia AJA, Bulloch S;
 XX WPI; 2003-457603/43.
 DR N-PSDB; AAL60550.
 XX The invention relates to human organelle-associated proteins (ORGA) and
 CC nucleic acid molecules encoding such proteins. ORGA sequences are useful
 CC for diagnosing, treating and/or preventing cell proliferative disorders
 CC (e.g. actinic keratosis, arteriosclerosis, bursitis, cirrhosis, cancers
 CC or hepatitis), reproductive disorders (e.g. infertility, endometriosis,
 CC ovulatory defects, ectopic pregnancy or teratogenesis), gastrointestinal
 CC disorders (e.g. dysphagia, peptide oesophagitis, gastritis, indigestion,
 CC anorexia or gastroenteritis), neurological disorders (e.g. Alzheimer's
 CC disease, Pick's disease, Huntington's disease or dementia), urologic
 CC disorders (e.g. infection, urethritis, cystitis, prostatitis or upper

DB 1 MQTCPLAFPGHVSQALGTLFLAASLSAQNEGWDSPICTEGVVSVSGENTVMSCNISNA 60
 QY 61 FSHVNIKLRAHQESAIENEVAPGYFSDGWLQVGGVAQLVIKGARDSHAGLYMHLV 120
 DB 61 FSHVNIKLRAHQESAIENEVAPGYFSDGWLQVGGVAQLVIKGARDSHAGLYMHLV 120
 QY 121 GHQ 123
 DB 121 GHQ 123
 RESULT 8
 ID AAO29897 standard; protein; 183 AA.
 XX AAO29897;
 AC AAO29897;
 DT 03-SEP-2003 (first entry)
 XX Human organelle-associated protein (ORGA)-10.
 DE Human; organelle-associated protein; ORGA; cell proliferative disorder;
 XX actinic keratosis; arteriosclerosis; reproductive disorder; infertility;
 KW endometriosis; gastrointestinal disorder; ectopic pregnancy; hepatitis;
 KW gastroenteritis; neurological disorder; Alzheimer's disease; cirrhosis;
 KW glomerulonephritis; renal amyloidosis; renal failure; Addison's disease;
 KW renal disorder; ovulatory defect; teratogenesis; Pick's disease; cancer;
 KW Huntington's disease; urologic disorder; cystitis; dementia; dysphagia;
 KW indigestion; gastritis; hypertension; gene therapy; bursitis; anorexia;
 KW infection; urethritis; prostatitis.
 XX Homo sapiens.
 OS WO2003044171-A2.
 PN 30-MAY-2003.
 PD 15-NOV-2002; 2002WO-US036807.
 XX 16-NOV-2001; 2001US-0332384P.
 PR 13-DEC-2001; 2001US-0341187P.
 PR 23-JAN-2002; 2002US-0351151P.
 PR 27-FEB-2002; 2002US-0360269P.
 PR 05-APR-2002; 2002US-0370637P.
 PR 14-JUN-2002; 2002US-0388946P.
 XX (INCY-) INCYTE GENOMICS INC.
 PA Rankumar J, Burford N, Xu Y, Chawla NK, Kable AE, Becha SD;
 PI Richardson TW, Gorvad AE, Sprague WW, Elliott VS, Khare R;
 PI Hawkins PR, Jin P, Lee SY, Marquis JP, Swarnakar A, Wilson AD;
 PI Hafalia AJA, Bulloch S;
 XX WPI; 2003-457603/43.
 DR N-PSDB; AAL60550.
 XX New organelle-associated proteins and polynucleotides, useful for
 CC diagnosing, treating and/or preventing cell proliferative, reproductive,
 CC gastrointestinal, neurological, urologic, and renal disorders.
 PS Claim 1; Page 166-167; 194pp; English.
 XX The invention relates to human organelle-associated proteins (ORGA) and
 CC nucleic acid molecules encoding such proteins. ORGA sequences are useful
 CC for diagnosing, treating and/or preventing cell proliferative disorders
 CC (e.g. actinic keratosis, arteriosclerosis, bursitis, cirrhosis, cancers
 CC or hepatitis), reproductive disorders (e.g. infertility, endometriosis,
 CC ovulatory defects, ectopic pregnancy or teratogenesis), gastrointestinal
 CC disorders (e.g. dysphagia, peptide oesophagitis, gastritis, indigestion,
 CC anorexia or gastroenteritis), neurological disorders (e.g. Alzheimer's
 CC disease, Pick's disease, Huntington's disease or dementia), urologic
 CC disorders (e.g. infection, urethritis, cystitis, prostatitis or upper

PS Claim 1; Page 165-166; 194pp; English.

XX The invention relates to human organelle-associated proteins (ORGA) and
CC nucleic acid molecules encoding such proteins. ORGA sequences are useful
CC for diagnosing, treating and/or preventing cell proliferative disorders
CC (e.g. actinic keratosis, arteriosclerosis, bursitis, cirrhosis, cancers
CC or hepatitis), reproductive disorders (e.g. infertility, endometriosis,
CC ovulatory defects, ectopic pregnancy or teratogenesis), gastrointestinal
CC disorders (e.g. dysphagia, peptide esophagitis, gastritis, indigestion,
CC anorexia or gastroenteritis), neurological disorders (e.g. Alzheimer's
CC disease, Pick's disease, Huntington's disease or dementia), urologic
CC disorders (e.g. infection, urethritis, cystitis, prostatitis or upper
CC tract infections) and renal disorders (e.g. glomerulonephritis, renal
CC amyloidosis, renal failure, Addison's disease or hypertension). The
CC invention is also useful in gene therapy. The present sequence is human
CC ORGA protein

XX Sequence 101 AA;

Query Match 37.1%; Score 92; DB 6; Length 101;
Best Local Similarity 100.0%; Pred. No. 1.1e-84;
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MOTCLAPFGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSNCISNA 60
Db 1 MOTCLAPFGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSNCISNA 60

Qy 61 FSHVNIKLRAHQESAIFNEVAPGVFSRDGQ 92

Db 61 FSHVNIKLRAHQESAIFNEVAPGVFSRDGQ 92

RESULT 10

AAB34671

ID AAB34671 standard; protein; 55 AA.

XX AAB34671;

XX 26-JAN-2001 (first entry)

XX Gene 35 human secreted protein homologous amino acid sequence #155.

XX Human; secreted protein; diagnosis; immunosuppressive; antiarthritic;
KW antiirrhematic; antiproliferative; cytostatic; cardiant; vasotropic;
KW cerebroprotective; neurotropic; neuroprotective; antibacterial; virucide;
KW fungicide; ophthalmological; gene therapy; autoimmune disease; neoplasm;
KW hyperproliferative disorder; cancer; cardiovascular disorder; infection;
KW cerebrovascular disorder; angiogenesis; nervous system disorder;
KW ocular disorder; wound healing; skin aging; food additive; preservative.

XX Homo sapiens.

XX WO200056751-A1.

XX 28-SEP-2000.

XX 09-MAR-2000; 2000WO-US006013.

XX 19-MAR-1999; 93US-0125360P.

XX 11-JUN-1999; 93US-0138626P.

XX 03-DEC-1999; 93US-0168662P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM, Komatsoulis G;

XX WPI; 2000-579482/54.

XX Isolated nucleic acid molecule encoding a human secreted protein is used
XX in preventing, treating or ameliorating a medical condition.

XX Disclosure; Page 412; 419pp; English.

XX

CC The polynucleotide sequences given in AAC59738 to AAC59787 encode the
CC human secreted proteins given in AAB34577 to AAB34626. AAB34627 to
CC AAB34686 represent human secreted polypeptide sequences and proteins
CC homologous to them, which are given in the exemplification of the present
CC invention. Human secreted proteins have activities based on the tissues
CC and cells the genes are expressed in. Example of activities include:
CC antiarthritic; immunosuppressive; antirrhematic; antiproliferative;
CC cytostatic; cardiant; vasotropic; cerebroprotective; neurotropic;
CC neuroprotective; antibacterial; virucide; fungicide; and
CC ophthalmological. The polynucleotides and proteins can be used to
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
CC in diagnosing a pathological condition or susceptibility to a
CC pathological condition. Disorders which are diagnosed or treated include
CC autoimmune diseases, hyperproliferative disorders e.g. neoplasms and
CC cancers of the breast or liver, cardiovascular disorders, cerebrovascular
CC disorders, angiogenesis, nervous system disorders, infections caused by
CC bacteria, viruses and fungi and ocular disorders. The proteins can also
CC be used to aid wound healing and epithelial cell proliferation, to
CC prevent skin aging due to sunburn, to maintain organs before
CC transplantation, for supporting cell culture of primary tissues, to
CC regenerate tissues and in chemotaxis. The proteins can also be used as a
CC food additive or preservative to increase or decrease storage
CC capabilities. AAC59729 to AAC59737 and AAB34576 represent sequences used
CC in the exemplification of the present invention

XX Sequence 55 AA;

Query Match 22.2%; Score 55; DB 3; Length 55;

Best Local Similarity 100.0%; Pred. No. 1.8e-47;

Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 194 LRAGAQQGLSRASAEELWTPDSEPTPRPLALVFKPSPLGALLSPQLPFPVAADP 248

Db 1 LRAGAQQGLSRASAEELWTPDSEPTPRPLALVFKPSPLGALLSPQLPFPVAADP 55

RESULT 11

AAO29898

ID AAO29898 standard; protein; 107 AA.

XX AAO29898;

XX 03-SEP-2003 (first entry)

XX Human organelle-associated protein (ORGA)-11.

XX Human; organelle-associated protein; ORGA; cell proliferative disorder;
KW actinic keratosis; arteriosclerosis; reproductive disorder; infertility;
KW endometriosis; gastrointestinal disorder; ectopic pregnancy; hepatitis;
KW gastroenteritis; neurological disorder; Alzheimer's disease; cirrhosis;
KW glomerulonephritis; renal amyloidosis; renal failure; Addison's disease;
KW renal disorder; ovulatory defect; teratogenesis; Pick's disease; cancer;
KW Huntington's disease; urologic disorder; cystitis; dementia; dysphagia;
KW indigestion; gastritis; hypertension; gene therapy; bursitis; anorexia;
KW infection; urethritis; prostatitis.

XX Homo sapiens.

XX WO2003044171-A2.

XX 30-MAY-2003.

XX 15-NOV-2002; 2002WO-US036807.

XX 16-NOV-2001; 2001US-0332384P.

XX 13-DEC-2001; 2001US-0341187P.

XX 23-JAN-2002; 2002US-0351151P.

XX 27-FEB-2002; 2002US-0360269P.

XX 05-APR-2002; 2002US-0370637P.

XX 14-JUN-2002; 2002US-0388946P.

XX (INCY-) INCYTE GENOMICS INC.

PI Rosen CA, Ruben SM, Komatsoulis G;
 XX WPI; 2000-579482/54.
 DR
 XX Isolated nucleic acid molecule encoding a human secreted protein is used
 PT in preventing, treating or ameliorating a medical condition.
 PT
 XX Disclosure; Page 412-413; 419pp; English.
 PS
 XX The polynucleotide sequences given in AAC59738 to AAC59787 encode the
 CC human secreted proteins given in AAB34577 to AAB34626. AAB34627 to
 CC AAB34686 represent human secreted polypeptide sequences and proteins
 CC homologous to them, which are given in the exemplification of the present
 CC invention. Human secreted proteins have activities based on the tissues
 CC and cells the genes are expressed in. Example of activities include:
 CC antiarthritic; immunosuppressive; antirheumatic; antiproliferative;
 CC cytotatic; cardiatic; vasotropic; cerebroprotective; neurotropic;
 CC neuroprotective; antibacterial; virucide; fungicide; and
 CC ophthalmological. The polynucleotides and proteins can be used to
 CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
 CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
 CC in diagnosing a pathological condition or susceptibility to a
 CC pathological condition. Disorders which are diagnosed or treated include
 CC autoimmune diseases, hyperproliferative disorders e.g. neoplasms and
 CC cancers of the breast or liver, cardiovascular disorders, cerebrovascular
 CC disorders, angiogenesis, nervous system disorders, infections caused by
 CC bacteria, viruses and fungi and ocular disorders. The proteins can also
 CC be used to aid wound healing and epithelial cell proliferation, to
 CC prevent skin aging due to sunburn, to maintain organs before
 CC transplantation, for supporting cell culture of primary tissues, to
 CC regenerate tissues and in chemotaxis. The proteins can also be used as a
 CC food additive or preservative to increase or decrease storage
 CC capabilities. AAC59729 to AAC59737 and AAB34576 represent sequences used
 CC in the exemplification of the present invention
 XX Sequence 55 AA;
 SQ
 Query Match 19.4%; Score 48; DB 3; Length 55;
 Best Local Similarity 100.0%; Pred. No. 2.2e-40;
 Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 194 LRAGNQQGLSRASAEIWTPTDSEPTPRPLALVFKPSPLGALLELSPQL 241
 DB 1 LRAGNQQGLSRASAEIWTPTDSEPTPRPLALVFKPSPLGALLELSPQL 48
 RESULT 13.
 ADB36337
 ID ADB36337 standard; protein; 43 AA.
 XX
 AC ADB36337;
 XX
 DT 04-DEC-2003 (first entry)
 DE Human immune response associated protein IRAP-8 SEQ ID NO:8.
 XX human; immune response associated protein; IRAP; anti-HIV; anti-allergic;
 KW antianemic; antiasthmatic; antiarteriosclerotic; dermatological;
 KW antinflammatory; antidiabetic; nephrotropic; antithyroid; thyromimetic;
 KW immunosuppressive; antirheumatic; antiarthritic; osteopathic;
 KW antibacterial; virucide; antiparasitic; protozoacide; fungicide;
 KW cerebroprotective; neuroprotective; neurotropic; antiparkinsonian;
 KW antipsoriatic; cytotatic; cardiatic; gene therapy;
 KW immune system disorder; neurological disorder; developmental disorder;
 KW muscle disorder; cell proliferative disorder; AIDS; allergy; anaemia;
 KW asthma; atherosclerosis; contact dermatitis; Crohn's disease; diabetes;
 KW glomerulonephritis; Grave's disease; Hashimoto's thyroiditis;
 KW multiple sclerosis; rheumatoid arthritis; osteoporosis;
 KW systemic lupus erythematosus; infection; stroke; Alzheimer's disease;
 KW Parkinson's disease; psoriasis; cancer; cardiomyopathy.
 OS Homo sapiens.
 XX
 XX
 PI Ramkumar J, Burford N, Xu Y, Chawla NK, Kable AE, Becha SD;
 PI Richardson TW, Gorvad AE, Sprague WW, Elliott VS, Khare R;
 PI Hawkins PR, Jin P, Lee SY, Marquis JP, Swarnakar A, Wilson AD;
 PI Hafalia AJA, Bulloch S;
 XX WPI; 2003-457603/43.
 DR N-PSDB; AAL60551.
 XX
 XX New organelle-associated proteins and polynucleotides, useful for
 PT diagnosing, treating and/or preventing cell proliferative, reproductive,
 PT gastrointestinal, neurological, urologic, and renal disorders.
 PT
 XX Claim 1; Page 167; 194pp; English.
 PS
 XX The invention relates to human organelle-associated proteins (ORGA) and
 CC nucleic acid molecules encoding such proteins. ORGA sequences are useful
 CC for diagnosing, treating and/or preventing cell proliferative disorders
 CC (e.g. actinic keratosis, arteriosclerosis, bursitis, cirrhosis, cancers
 CC or hepatitis), reproductive disorders (e.g. infertility, endometriosis,
 CC ovulatory defects, ectopic pregnancy or teratogenesis), gastrointestinal
 CC disorders (e.g. dysphagia, peptide oesophagitis, gastritis, indigestion,
 CC anorexia or gastroenteritis), neurological disorders (e.g. Alzheimer's
 CC disease, Pick's disease, Huntington's disease or dementia), urologic
 CC disorders (e.g. infection, urethritis, cystitis, prostatitis or upper
 CC tract infections) and renal disorders (e.g. glomerulonephritis, renal
 CC amyloidosis, renal failure, Addison's disease or hypertension). The
 CC invention is also useful in gene therapy. The present sequence is human
 CC ORGA protein.
 XX Sequence 107 AA;
 SQ
 Query Match 22.2%; Score 55; DB 6; Length 107;
 Best Local Similarity 100.0%; Pred. No. 3.4e-47;
 Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MQTCPLAFCHVSOALGTLFLAASLSAONEGWDSPTCTGVVSVSGWNTWMS 55
 DB 1 MQTCPLAFCHVSOALGTLFLAASLSAONEGWDSPTCTGVVSVSGWNTWMS 55
 RESULT 12
 AAB34672
 ID AAB34672 standard; protein; 55 AA.
 XX
 AC AAB34672;
 XX
 DT 26-JAN-2001 (first entry)
 DE Human secreted protein sequence encoded by gene 35 SEQ ID NO:156.
 XX
 KW Human; secreted protein; diagnosis; immunosuppressive; antiarthritic;
 KW antirheumatic; antiproliferative; cytotatic; cardiatic; vasotropic;
 KW cerebroprotective; neurotropic; neuroprotective; antibacterial; virucide;
 KW fungicide; ophthalmological; gene therapy; autoimmune disease; neoplasm;
 KW hyperproliferative disorder; cancer; cardiovascular disorder; infection;
 KW cerebrovascular disorder; angiogenesis; nervous system disorder;
 KW ocular disorder; wound healing; skin aging; food additive; preservative.
 OS Homo sapiens.
 XX
 XX WO200056751-A1.
 PN
 XX 28-SEP-2000.
 PD
 XX 09-MAR-2000; 2000WO-US006013.
 PF
 XX 19-MAR-1999; 99US-0125360P.
 PR 11-JUN-1999; 99US-0138626P.
 PR 03-DEC-1999; 99US-0168662P.
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX

PN WO2003074726-A2.
XX 12-SEP-2003.
XX 28-FEB-2003; 2003WO-US006307.
XX 01-MAR-2002; 2002US-0361088P.
PR 27-MAR-2002; 2002US-0368494P.
PR 10-MAY-2002; 2002US-0379876P.
PR 28-JUN-2002; 2002US-0392641P.
XX (INCY-) INCYTE GENOMICS INC.
XX Chawla NK, Becha SD, Jin P, Tran UK, Kable AE, Lee S;
PI Marquis JP, Emerling BM, Ison CH, Ramkumar J, Chang H;
XX WPI; 2003-722079/58.
DR N-PSDB; ADB36353.
XX New human immune response associated proteins and polynucleotides, useful
PT for diagnosing, preventing or treating diseases or conditions associated
PT with aberrant enzyme expression, e.g. cancer, AIDS, atherosclerosis or
PT stroke.
XX Claim 1; Page 143; 158pp; English.
XX ADB36346 to ADB36361 encode the human immune response associated proteins
CC given in ADB36330 to ADB36345, designated IRAP-1 to IRAP-16. IRAP
CC sequences have anti-HIV, anti-allergic, anti-naemic, antiasthmatic,
CC antiarteriosclerotic, dermatological, anti-inflammatory, antidiabetic,
CC nephrotropic, antithyroid, thyromimetic, immunosuppressive,
CC antirheumatic, antiarthritic, osteoprotective, antibacterial, virucide,
CC antiparasitic, protozoacide, fungicide, cerebroprotective,
CC neuroprotective, neurotropic, antiparkinsonian, antipsoriatic, cytostatic
CC and cardiant activities, and can be used in gene therapy. The IRAP
CC proteins and polynucleotides can be used in diagnosing, preventing or
CC treating diseases or conditions associated with the decreased expression
CC or overexpression of IRAP, such as immune system, neurological,
CC developmental, muscle or cell proliferative disorders. The disorders may
CC include AIDS, allergies, anaemia, asthma, atherosclerosis, contact
CC dermatitis, Crohn's disease, diabetes, glomerulonephritis, Grave's
CC disease, Hashimoto's thyroiditis, multiple sclerosis, rheumatoid
CC arthritis, osteoporosis, systemic lupus erythematosus, infections (e.g.
CC bacterial, viral, fungal, parasitic or protozoal), stroke, Alzheimer's
CC disease, Parkinson's disease, psoriasis, cancer and cardiomyopathy. They
CC are also useful in assessing the effects of exogenous compounds on the
CC expression of nucleic acid and amino acid sequences of IRAP. The IRAP or
CC its fragments are useful in screening compounds for effectiveness as
CC agonist or antagonist of the polypeptides, or in altering the expression
CC of the target polynucleotide and compounds that specifically bind to or
CC modulate the activity of the polypeptide.
XX Sequence 43 AA;
SQ Query Match 12.5%; Score 31; DB 7; Length 43;
Best Local Similarity 100.0%; Pred. No. 2.9e-23;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MOTCLAPFGHVSQALGTLFLAASLSAQNE 31
DB 1 MOTCLAPFGHVSQALGTLFLAASLSAQNE 31
RESULT 14
AAB34673
ID AAB34673 standard; protein; 30 AA.
XX AAB34673;
XX AC
XX 26-JAN-2001 (first entry)
XX DT
XX Gene 35 human secreted protein homologous amino acid sequence #157.
XX DE
XX AC

KW Human; secreted protein; diagnosis; immunosuppressive; antiarthritic;
KW antirheumatic; antiproliferative; cytostatic; cardiac; vasotropic;
KW cerebroprotective; neurotropic; neuroprotective; antibacterial; virucide;
KW fungicide; ophthalmological; gene therapy; autoimmune disease; neoplasm;
KW hyperproliferative disorder; cancer; cardiovascular disorder; infection;
KW cerebrovascular disorder; angiogenesis; nervous system disorder;
KW ocular disorder; wound healing; skin aging; food additive; preservative.
OS Homo sapiens.
XX WO2000056751-A1.
PN 28-SEP-2000.
XX 09-MAR-2000; 2000WO-US006013.
XX 19-MAR-1999; 99US-0125360P.
PR 11-JUN-1999; 99US-0138626P.
PR 03-DEC-1999; 99US-0168626P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Ruben SM, Komatsoulis G;
PI WPI; 2000-579482/54.
XX Isolated nucleic acid molecule encoding a human secreted protein is used
PT in preventing, treating or ameliorating a medical condition.
XX Disclosure; Page 413; 419pp; English.
XX The polynucleotide sequences given in AAC59738 to AAC59787 encode the
CC human secreted proteins given in AAB34577 to AAB34626. AAB34627 to
CC AAB34686 represent human secreted polypeptide sequences and proteins
CC homologous to them, which are given in the exemplification of the present
CC invention. Human secreted proteins have activities based on the tissues
CC and cells the genes are expressed in. Example of activities include:
CC antiarthritic; immunosuppressive; antirheumatic; antiproliferative;
CC cytostatic; cardiac; vasotropic; cerebroprotective; neurotropic;
CC neuroprotective; antibacterial; virucide; fungicide; and
CC ophthalmological. The polynucleotides and proteins can be used to
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
CC in diagnosing a pathological condition or susceptibility to a
CC pathological condition. Disorders which are diagnosed or treated include
CC autoimmune diseases, hyperproliferative disorders e.g. neoplasms and
CC cancers of the breast or liver, cardiovascular disorders, cerebrovascular
CC disorders, angiogenesis, nervous system disorders, infections caused by
CC bacteria, viruses and fungi and ocular disorders. The proteins can also
CC be used to aid wound healing and epithelial cell proliferation, to
CC prevent skin aging due to sunburn, to maintain organs before
CC transplantation, for supporting cell culture of primary tissues, to
CC regenerate tissues and in chemotaxis. The proteins can also be used as a
CC food additive or preservative to increase or decrease storage
CC capabilities. AAC59729 to AAC59737 and AAB34576 represent sequences used
CC in the exemplification of the present invention
XX Sequence 30 AA;
SQ Query Match 12.1%; Score 30; DB 3; Length 30;
Best Local Similarity 100.0%; Pred. No. 2.1e-22;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 62 SHVNIKLRAGQESAIENFVAPGYFSRDGW 91
DB 1 SHVNIKLRAGQESAIENFVAPGYFSRDGW 30
.. RESULT 15
AAB34674
ID AAB34674 standard; protein; 27 AA.
XX AAB34674;
XX AC

Search completed: March 20, 2005, 10:16:44
Job time : 91 secs

XX DT 26-JAN-2001 (first entry)
XX DE Human secreted protein sequence encoded by gene 35 SEQ ID NO:158.
XX KW Human; secreted protein; diagnosis; immunosuppressive; antiarthritic;
KW antirheumatic; antiproliferative; cytostatic; cardiant; vasotropic;
KW cerebroprotective; neurotropic; neuroprotective; antibacterial; virucide;
KW fungicide; ophthalmological; gene therapy; autoimmune disease; neoplasm;
KW hyperproliferative disorder; cancer; cardiovascular disorder; infection;
KW cerebrovascular disorder; angiogenesis; nervous system disorder;
KW ocular disorder; wound healing; skin aging; food additive; preservative.
XX OS Homo sapiens.
XX PN WO200056751-A1.
XX PD 28-SEP-2000.
XX PF 09-MAR-2000; 2000WO-US006013.
XX PR 19-MAR-1999; 99US-0125360P.
PR 11-JUN-1999; 99US-0138626P.
PR 03-DEC-1999; 99US-0168662P.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PI Rosen CA, Ruben SM, Komatsoulis G;
XX WPI; 2000-579482/54.
XX PT Isolated nucleic acid molecule encoding a human secreted protein is used
in preventing, treating or ameliorating a medical condition.
XX PS Disclosure; Page 413; 419pp; English.
XX CC The polynucleotide sequences given in AAC59738 to AAC59787 encode the
human secreted proteins given in AAB34577 to AAB34626. AAB34627 to
AAB34686 represent human secreted polypeptide sequences and proteins
homologous to them, which are given in the exemplification of the present
invention. Human secreted proteins have activities based on the tissues
and cells the genes are expressed in. Example of activities include:
CC antiarthritic; immunosuppressive; antirheumatic; antiproliferative;
CC cytostatic; cardiant; vasotropic; cerebroprotective; neurotropic;
CC neuroprotective; antibacterial; virucide; fungicide; and
CC ophthalmological. The polynucleotides and proteins can be used to
prevent, treat or ameliorate a medical condition in e.g. humans, mice,
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
in diagnosing a pathological condition or susceptibility to a
CC pathological condition. Disorders which are diagnosed or treated include
CC autoimmune diseases, hyperproliferative disorders e.g. neoplasms and
CC cancers of the breast or liver, cardiovascular disorders, cerebrovascular
CC disorders, angiogenesis, nervous system disorders, infections caused by
CC bacteria, viruses and fungi and ocular disorders. The proteins can also
be used to aid wound healing and epithelial cell proliferation, to
CC prevent skin aging due to sunburn, to maintain organs before
CC transplantation, for supporting cell culture of primary tissues, to
CC regenerate tissues and in chemotaxis. The proteins can also be used as a
CC food additive or preservative to increase or decrease storage
CC capabilities. AAC59729 to AAC59737 and AAB34576 represent sequences used
in the exemplification of the present invention
XX SQ Sequence 27 AA;

Query Match 4.8%; Score 12; DB 3; Length 27;
Best Local Similarity 100.0%; Pred. No. 0.00033;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 SHVNIKLRAHQ 73
Db 1 SHVNIKLRAHQ 12

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 20, 2005, 10:13:54 ; Search time 26 Seconds
(without alignments)
712.037 Million cell updates/sec

Title: US-10-080-522-1
Perfect score: 248
Sequence: 1 MOTCPAPFGHVSQLGTL.....FLGALELLSQPLFPYADP 248

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 513545 seqs, 74649064 residues

Word size : 0
Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfiles.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	248	100.0	248	US-08-755-559-1	Sequence 1, Appli
2	248	100.0	248	US-09-210-474-1	Sequence 1, Appli
3	248	100.0	248	US-09-539-774-1	Sequence 1, Appli
4	248	100.0	248	US-09-997-165-4	Sequence 4, Appli
5	248	100.0	248	US-09-949-016-6297	Sequence 6297, Ap
6	248	100.0	258	US-09-949-016-8807	Sequence 8807, Ap
7	8	3.2	453	US-09-252-991A-31634	Sequence 31634, A
8	7	2.8	67	US-09-248-796A-23640	Sequence 23640, A
9	7	2.8	81	US-09-621-976-5355	Sequence 5255, Ap
10	7	2.8	83	US-09-270-767-60208	Sequence 60208, A
11	7	2.8	93	US-09-902-540-13818	Sequence 13818, A
12	7	2.8	138	US-09-148-545-222	Sequence 222, App
13	7	2.8	139	US-09-148-545-160	Sequence 160, App
14	7	2.8	197	US-09-252-991A-32189	Sequence 32189, A
15	7	2.8	202	US-09-902-540-16034	Sequence 16034, A
16	7	2.8	210	US-10-006-011A-10	Sequence 10, Appl
17	7	2.8	231	US-09-583-110-3995	Sequence 3995, Ap
18	7	2.8	233	US-09-252-991A-21274	Sequence 21274, A
19	7	2.8	235	US-09-390-721-4	Sequence 4, Appli
20	7	2.8	235	US-09-489-039A-12572	Sequence 12572, A
21	7	2.8	235	US-09-713-893-4	Sequence 4, Appli
22	7	2.8	242	US-08-663-310-9	Sequence 9, Appli
23	7	2.8	242	US-09-006-491-9	Sequence 9, Appli
24	7	2.8	242	US-09-135-919-9	Sequence 9, Appli
25	7	2.8	242	US-08-980-832-29	Sequence 29, Appl
26	7	2.8	242	US-09-920-923B-29	Sequence 29, Appl
27	7	2.8	248	US-09-050-739-60	Sequence 60, Appl

28	7	2.8	262	4	US-09-107-433-2779	Sequence 2779, Ap
29	7	2.8	278	5	PCT-US96-03916-60	Sequence 60, Appl
30	7	2.8	278	5	PCT-US96-03916-72	Sequence 72, Appl
31	7	2.8	284	4	US-10-006-011A-9	Sequence 9, Appli
32	7	2.8	286	4	US-09-543-681A-5923	Sequence 5923, Ap
33	7	2.8	325	4	US-09-270-767-44750	Sequence 44750, A
34	7	2.8	329	4	US-09-724-797-6	Sequence 6, Appli
35	7	2.8	336	4	US-09-489-039A-8430	Sequence 8430, Ap
36	7	2.8	345	4	US-09-602-787A-544	Sequence 544, App
37	7	2.8	353	4	US-09-902-540-14163	Sequence 14163, A
38	7	2.8	358	4	US-09-328-352-6375	Sequence 6375, Ap
39	7	2.8	361	4	US-09-668-097A-36	Sequence 36, Appl
40	7	2.8	362	4	US-09-107-433-3284	Sequence 3284, Ap
41	7	2.8	399	4	US-09-134-000C-5890	Sequence 5890, Ap
42	7	2.8	429	4	US-09-270-767-43353	Sequence 43353, A
43	7	2.8	436	4	US-09-252-991A-19792	Sequence 19792, A
44	7	2.8	506	4	US-09-252-991A-30561	Sequence 30561, A
45	7	2.8	509	3	US-08-369-822C-26	Sequence 26, Appl

ALIGNMENTS

RESULT 1
US-08-755-559-1
; Sequence 1, Application US/08755559
; Patent No. 5912142
; GENERAL INFORMATION:
; APPLICANT: KAUFMAN, RUSSEL E.
; APPLICANT: SLENTZ-KESLER, KIMBERLY
; TITLE OF INVENTION: GENE PRODUCT OVER EXPRESSED IN CANCER
; TITLE OF INVENTION: CELLS
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHVE P.C.
; STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/755,559
; FILING DATE: 22-NOV-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: WILSON, MARY J.
; REGISTRATION NUMBER: 32,955
; REFERENCE/DOCKET NUMBER: 1579-116
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 248 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-755-559-1

Query Match 100.0%; Score 248; DB 2; Length 248;
Best Local Similarity 100.0%; Pred. No. 9.8e-238;
Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MOTCPAPFGHVSQLGTLFLAASLSAQNKGWDSPTCTEGVSVSWGENTVMSCNISA 60
DB 1 MOTCPAPFGHVSQLGTLFLAASLSAQNKGWDSPTCTEGVSVSWGENTVMSCNISA 60

QY 61 FSHVNIKLRAHQESAI FNEVAPGYFSRDGWLQVQGGVAQLVIKGARDSHAGLYMHLV 120
 DB 61 FSHVNIKLRAHQESAI FNEVAPGYFSRDGWLQVQGGVAQLVIKGARDSHAGLYMHLV 120
 QY 121 GHORNNRQVTLVSGAEPOSAPDTGFWPVPVAVTAVFILLVALVNFAMVYRCRCQORREK 180
 DB 121 GHORNNRQVTLVSGAEPOSAPDTGFWPVPVAVTAVFILLVALVNFAMVYRCRCQORREK 180
 QY 181 KFFLEPQMKVAALRAGAQOGLSRASAEIWTDPSEPTPRPLALVFKPSPLGALELLSPQP 240
 DB 181 KFFLEPQMKVAALRAGAQOGLSRASAEIWTDPSEPTPRPLALVFKPSPLGALELLSPQP 240
 QY 241 LFPYAADP 248
 DB 241 LFPYAADP 248

RESULT 2
 US-09-210-474-1
 ; Sequence 1, Application US/09210474
 ; Patent No. 6072034
 ; GENERAL INFORMATION:
 ; APPLICANT: KAUFMAN, RUSSEL E.
 ; APPLICANT: SLENTZ-KESLER, KIMBERLY
 ; TITLE OF INVENTION: GENE PRODUCT OVER EXPRESSED IN CANCER
 ; TITLE OF INVENTION: CELLS
 ; NUMBER OF SEQUENCES: 2
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: NIXON & VANDERHUYE P.C.
 ; STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
 ; CITY: ARLINGTON
 ; STATE: VIRGINIA
 ; COUNTRY: U.S.A.
 ; ZIP: 22201-4714
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/210,474
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/755,559
 ; FILING DATE: 22-NOV-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: WILSON, MARY J.
 ; REGISTRATION NUMBER: 32,955
 ; REFERENCE/DOCKET NUMBER: 1579-116
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (703) 816-4000
 ; TELEFAX: (703) 816-4100
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 248 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS:
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-09-210-474-1

Query Match 100.0%; Score 248; DB 3; Length 248;
 Best Local Similarity 100.0%; Pred. No. 9.8e-238;
 Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MOTCLAPFGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSCNISNA 60
 DB 1 MOTCLAPFGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSCNISNA 60
 QY 61 FSHVNIKLRAHQESAI FNEVAPGYFSRDGWLQVQGGVAQLVIKGARDSHAGLYMHLV 120
 DB 61 FSHVNIKLRAHQESAI FNEVAPGYFSRDGWLQVQGGVAQLVIKGARDSHAGLYMHLV 120

QY 121 GHORNNRQVTLVSGAEPOSAPDTGFWPVPVAVTAVFILLVALVNFAMVYRCRCQORREK 180
 DB 121 GHORNNRQVTLVSGAEPOSAPDTGFWPVPVAVTAVFILLVALVNFAMVYRCRCQORREK 180
 QY 181 KFFLEPQMKVAALRAGAQOGLSRASAEIWTDPSEPTPRPLALVFKPSPLGALELLSPQP 240
 DB 181 KFFLEPQMKVAALRAGAQOGLSRASAEIWTDPSEPTPRPLALVFKPSPLGALELLSPQP 240
 QY 241 LFPYAADP 248
 DB 241 LFPYAADP 248

RESULT 3
 US-09-539-774-1
 ; Sequence 1, Application US/09539774
 ; Patent No. 6350615
 ; GENERAL INFORMATION:
 ; APPLICANT: KAUFMAN, RUSSEL E.
 ; APPLICANT: SLENTZ-KESLER, KIMBERLY
 ; TITLE OF INVENTION: GENE PRODUCT OVER EXPRESSED IN CANCER
 ; TITLE OF INVENTION: CELLS
 ; NUMBER OF SEQUENCES: 2
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: NIXON & VANDERHUYE P.C.
 ; STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
 ; CITY: ARLINGTON
 ; STATE: VIRGINIA
 ; COUNTRY: U.S.A.
 ; ZIP: 22201-4714
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/539,774
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 09/210,474
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: WILSON, MARY J.
 ; REGISTRATION NUMBER: 32,955
 ; REFERENCE/DOCKET NUMBER: 1579-116
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (703) 816-4000
 ; TELEFAX: (703) 816-4100
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 248 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS:
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-09-539-774-1

Query Match 100.0%; Score 248; DB 3; Length 248;
 Best Local Similarity 100.0%; Pred. No. 9.8e-238;
 Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MOTCLAPFGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSCNISNA 60
 DB 1 MOTCLAPFGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSCNISNA 60
 QY 61 FSHVNIKLRAHQESAI FNEVAPGYFSRDGWLQVQGGVAQLVIKGARDSHAGLYMHLV 120
 DB 61 FSHVNIKLRAHQESAI FNEVAPGYFSRDGWLQVQGGVAQLVIKGARDSHAGLYMHLV 120
 QY 121 GHORNNRQVTLVSGAEPOSAPDTGFWPVPVAVTAVFILLVALVNFAMVYRCRCQORREK 180
 DB 121 GHORNNRQVTLVSGAEPOSAPDTGFWPVPVAVTAVFILLVALVNFAMVYRCRCQORREK 180

Db 121 GHQNNRQVTLVSGAEPSQAPDTGFWPVPVAVTAVFILLVALVWFAYRCRCQORREK 180
Qy 181 KFFLLEPQMKVAALRAGAQOGLSRASAEIWTDPDSEPTPRPLALVFKPSPLGALLELLSQP 240
Db 181 KFFLLEPQMKVAALRAGAQOGLSRASAEIWTDPDSEPTPRPLALVFKPSPLGALLELLSQP 240
Qy 241 LFPYAADP 248
Db 241 LFPYAADP 248

RESULT 4

US-09-997-165-4
; Sequence 4, Application US/09997165
; Patent No. 6762030
; GENERAL INFORMATION:
; APPLICANT: Lyman, Stewart D.
; APPLICANT: Fanslow, William C.
; TITLE OF INVENTION: LIGAND FOR CD7 AND METHODS OF USE THEREOF
; FILE REFERENCE: 2913-US
; CURRENT APPLICATION NUMBER: US/09/997,165
; CURRENT FILING DATE: 2001-11-27
; PRIOR APPLICATION NUMBER: PCT/US00/14612
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/136,450
; PRIOR FILING DATE: 1999-05-28
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 248
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-997-165-4

Query Match 100.0%; Score 248; DB 4; Length 248;

Best Local Similarity 100.0%; Pred. No. 9.8e-238; Mismatches 0; Indels 0; Gaps 0;
Matches 248; Conservative 0;

Qy 1 MOTCLAPFGHVHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSGENTVMSCNISNA 60
Db 1 MOTCLAPFGHVHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSGENTVMSCNISNA 60
Qy 61 FSHVNIKLRAHQESAI FNEVAPGYFSRDGHWQVQGGVAQLVIKGARDSHAGLYMMHLV 120
Db 61 FSHVNIKLRAHQESAI FNEVAPGYFSRDGHWQVQGGVAQLVIKGARDSHAGLYMMHLV 120
Qy 121 GHQNNRQVTLVSGAEPSQAPDTGFWPVPVAVTAVFILLVALVWFAYRCRCQORREK 180
Db 121 GHQNNRQVTLVSGAEPSQAPDTGFWPVPVAVTAVFILLVALVWFAYRCRCQORREK 180
Qy 181 KFFLLEPQMKVAALRAGAQOGLSRASAEIWTDPDSEPTPRPLALVFKPSPLGALLELLSQP 240
Db 181 KFFLLEPQMKVAALRAGAQOGLSRASAEIWTDPDSEPTPRPLALVFKPSPLGALLELLSQP 240
Qy 241 LFPYAADP 248
Db 241 LFPYAADP 248

RESULT 5

US-09-949-016-6297
; Sequence 6297, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6297
; LENGTH: 248
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6297

Query Match 100.0%; Score 248; DB 4; Length 248;

Best Local Similarity 100.0%; Pred. No. 9.8e-238; Mismatches 0; Indels 0; Gaps 0;
Matches 248; Conservative 0;

Qy 1 MOTCLAPFGHVHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSGENTVMSCNISNA 60
Db 1 MOTCLAPFGHVHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSGENTVMSCNISNA 60
Qy 61 FSHVNIKLRAHQESAI FNEVAPGYFSRDGHWQVQGGVAQLVIKGARDSHAGLYMMHLV 120
Db 61 FSHVNIKLRAHQESAI FNEVAPGYFSRDGHWQVQGGVAQLVIKGARDSHAGLYMMHLV 120
Qy 121 GHQNNRQVTLVSGAEPSQAPDTGFWPVPVAVTAVFILLVALVWFAYRCRCQORREK 180
Db 121 GHQNNRQVTLVSGAEPSQAPDTGFWPVPVAVTAVFILLVALVWFAYRCRCQORREK 180
Qy 181 KFFLLEPQMKVAALRAGAQOGLSRASAEIWTDPDSEPTPRPLALVFKPSPLGALLELLSQP 240
Db 181 KFFLLEPQMKVAALRAGAQOGLSRASAEIWTDPDSEPTPRPLALVFKPSPLGALLELLSQP 240
Qy 241 LFPYAADP 248
Db 241 LFPYAADP 248

RESULT 6

US-09-949-016-8807
; Sequence 8807, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8807
; LENGTH: 258
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-8807

Query Match 100.0%; Score 248; DB 4; Length 258;

Best Local Similarity 100.0%; Pred. No. 1e-237; Mismatches 0; Indels 0; Gaps 0;
Matches 248; Conservative 0;

Qy 1 MOTCLAPFGHVHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSGENTVMSCNISNA 60
Db 11 MOTCLAPFGHVHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSGENTVMSCNISNA 70
Qy 61 FSHVNIKLRAHQESAI FNEVAPGYFSRDGHWQVQGGVAQLVIKGARDSHAGLYMMHLV 120
Db 71 FSHVNIKLRAHQESAI FNEVAPGYFSRDGHWQVQGGVAQLVIKGARDSHAGLYMMHLV 130

Sun Mar 20 14:17:41 2005

QY 121 GHORNRQVTLVSGAEQSPADTGFVPAVTVAVFILLVALVWFAYRCRCSCQRRREK 180
 Db 131 GHORNRQVTLVSGAEQSPADTGFVPAVTVAVFILLVALVWFAYRCRCSCQRRREK 190
 QY 181 KFFLEPQKVAALRAGAQGLSRASAEIWTDPSEPTPRPLALVFKPSPLGALLSPQP 240
 Db 191 KFFLEPQKVAALRAGAQGLSRASAEIWTDPSEPTPRPLALVFKPSPLGALLSPQP 250
 QY 241 LFPYAADP 248
 Db 251 LFPYAADP 258

RESULT 7
 US-09-252-991A-31634
 ; Sequence 31634, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; PRIOR FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 31634
 ; LENGTH: 453
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-31634

Query Match 3.2%; Score 8; DB 4; Length 453;
 Best Local Similarity 100.0%; Pred. No. 16;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 200 QGLSRASA 207
 Db 419 QGLSRASA 426

RESULT 8
 US-09-248-796A-23640
 ; Sequence 23640, Application US/09248796A
 ; Patent No. 6747137
 ; GENERAL INFORMATION:
 ; APPLICANT: Keith Weinstock et al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
 ; FILE REFERENCE: 107196.132
 ; CURRENT APPLICATION NUMBER: US/09/248,796A
 ; CURRENT FILING DATE: 1999-02-12
 ; PRIOR APPLICATION NUMBER: US 60/074,725
 ; PRIOR FILING DATE: 1998-02-13
 ; PRIOR APPLICATION NUMBER: US 60/096,409
 ; PRIOR FILING DATE: 1998-08-13
 ; NUMBER OF SEQ ID NOS: 28208
 ; SEQ ID NO 23640
 ; LENGTH: 67
 ; TYPE: PRT
 ; ORGANISM: Candida albicans
 US-09-248-796A-23640

Query Match 2.8%; Score 7; DB 4; Length 67;
 Best Local Similarity 100.0%; Pred. No. 26;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 175 QORREKK 181
 Db 37 QORREKK 43

RESULT 9
 US-09-621-976-5255
 ; Sequence 5255, Application US/09621976
 ; Patent No. 6639063
 ; GENERAL INFORMATION:
 ; APPLICANT: Dumas Milne Edwards, J.B.
 ; APPLICANT: Jobert, S.
 ; APPLICANT: Giordano, J.Y.
 ; TITLE OF INVENTION: ESTS and Encoded Human Proteins.
 ; FILE REFERENCE: GENSET.054PR2
 ; CURRENT APPLICATION NUMBER: US/09/621,976
 ; CURRENT FILING DATE: 2000-07-21
 ; NUMBER OF SEQ ID NOS: 19335
 ; SOFTWARE: Patent.pm
 ; SEQ ID NO 5255
 ; LENGTH: 81
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: SIGNAL
 ; LOCATION: -15..-1
 ; NAME/KEY: UNSURE
 ; LOCATION: 62
 ; OTHER INFORMATION: Xaa = *, Gln
 ; NAME/KEY: UNSURE
 ; LOCATION: 54
 ; OTHER INFORMATION: Xaa = Asp, Phe, Ile, Asn, Val, Tyr
 ; NAME/KEY: UNSURE
 ; LOCATION: 10
 ; OTHER INFORMATION: Xaa = Asp, Tyr
 ; NAME/KEY: UNSURE
 ; LOCATION: 14
 ; OTHER INFORMATION: Xaa = Cys, Arg
 ; NAME/KEY: UNSURE
 ; LOCATION: 64
 ; OTHER INFORMATION: Xaa = Cys, Gly
 ; NAME/KEY: UNSURE
 ; LOCATION: 38
 ; OTHER INFORMATION: Xaa = His, Tyr
 ; NAME/KEY: UNSURE
 ; LOCATION: 52
 ; OTHER INFORMATION: Xaa = Lys, Gln
 ; NAME/KEY: UNSURE
 ; LOCATION: 11
 ; OTHER INFORMATION: Xaa = Phe, Ile
 US-09-621-976-5255

Query Match 2.8%; Score 7; DB 4; Length 81;
 Best Local Similarity 100.0%; Pred. No. 31;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 TLLFLAA 24
 Db 5 TLLFLAA 11

RESULT 10
 US-09-270-767-60208
 ; Sequence 60208, Application US/09270767
 ; Patent No. 6703491
 ; GENERAL INFORMATION:
 ; APPLICANT: Homburger et al.
 ; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
 ; FILE REFERENCE: File Reference: 7326-094
 ; CURRENT APPLICATION NUMBER: US/09/270,767
 ; CURRENT FILING DATE: 1999-03-17
 ; NUMBER OF SEQ ID NOS: 62517
 ; SOFTWARE: Patentin Ver. 2.0
 ; SEQ ID NO 60208
 ; LENGTH: 83
 ; TYPE: PRT

```
; ORGANISM: Drosophila melanogaster
US-09-270-767-60208

Query Match          2.8%; Score 7; DB 4; Length 83;
Best Local Similarity 100.08; Pred. No. 32;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 149 VPAVVTA 155
Db 64 VPAVVTA 70

RESULT 11
US-09-902-540-13818
; Sequence 13818, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 13818
; LENGTH: 93
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-13818

Query Match          2.8%; Score 7; DB 4; Length 93;
Best Local Similarity 100.08; Pred. No. 36;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 25 SLSAQNE 31
Db 48 SLSAQNE 54

RESULT 12
US-09-148-545-222
; Sequence 222, Application US/09148545
; Patent No. 6590075
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 70 Human Secreted Proteins
; FILE REFERENCE: P2001P1
; CURRENT APPLICATION NUMBER: US/09/148,545
; CURRENT FILING DATE: 1998-09-04
; EARLIER APPLICATION NUMBER: PCT/US98/04482
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,161
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,626
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,334
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,336
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,163
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/047,615
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,600
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,597
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,502
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,633
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,583
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,617
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,618
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,503
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,592
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,581
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,584
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,500
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,587
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,492
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,598
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,613
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,582
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,596
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,612
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,632
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,601
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,580
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,568
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,314
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,569
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,311
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,671
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,674
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,669
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,312
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,313
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,672
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,315
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/056,886
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,877
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,889
```

us-10-080-522-1.oli.ra1

Sun Mar 20 14:17:41 2005

EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,893
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,630
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,878
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,662
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,872
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,882
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,637
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,903
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,888
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,879
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,880
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,894
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,911
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,636
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,874
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,845
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,892
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/047,595
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/057,761
EARLIER FILING DATE: 05-Sep-1997
EARLIER APPLICATION NUMBER: 60/047,599
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,588
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,585
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,586
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,590
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,594
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,589
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,593
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,614
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,578
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,576
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/047,501
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,670
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/056,632
EARLIER FILING DATE: 1997-08-22

EARLIER APPLICATION NUMBER: 60/056,664
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,876
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,881
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,909
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,875
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,862
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,887
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,908
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/048,964
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/057,650
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/056,884
EARLIER FILING DATE: 1997-08-22
NUMBER OF SEQ ID NOS: 280
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 222
LENGTH: 138

Query Match 2.8% Score 7; DB 4; Length 138;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 18 TLLFLAA 24
Db 5 TLLFLAA 11
|||||

RESULT 13
US-09-148-545-160
Sequence 160, Application US/09148545
Patent No. 6590075
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 70 Human Secreted Proteins
FILE REFERENCE: PZ001P1
CURRENT APPLICATION NUMBER: US/09/148,545
CURRENT FILING DATE: 1998-09-04
EARLIER APPLICATION NUMBER: PCT/US98/04482
EARLIER FILING DATE: 1998-03-06
EARLIER APPLICATION NUMBER: 60/040,162
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,333
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/038,621
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,161
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,626
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,334
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,336
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,163
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/047,615
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,600
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,597
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,502
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,633

[illegible]

EARLIER APPLICATION NUMBER: 60/056,666
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,872
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,882
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,637
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,903
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,888
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,879
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,880
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,894
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,874
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,910
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,864
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,631
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/047,595
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/057,761
EARLIER FILING DATE: 05-Sep-1997
EARLIER APPLICATION NUMBER: 60/047,599
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,588
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,585
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,586
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,590
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,594
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,589
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,593
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,614
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,578
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,576
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/047,501
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,670
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/056,632
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,664
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,876
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,881
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,909

```
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,875
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,862
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,887
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,908
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/057,650
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/056,884
; EARLIER FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 280
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 160
; LENGTH: 139

Query Match      2.8%; Score 7; DB 4; Length 139;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 TLLFLAA 24
Db 5 TLLFLAA 11

RESULT 14
US-09-252-991A-32189
; Sequence 32189, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196, 136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32189
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32189

Query Match      2.8%; Score 7; DB 4; Length 197;
Best Local Similarity 100.0%; Pred. No. 73;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 201 GLSRASA 207
Db 133 GLSRASA 139

RESULT 15
US-09-902-540-16034
; Sequence 16034, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
```

```
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 16034
; LENGTH: 202
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-16034
```

```
Query Match      2.8%; Score 7; DB 4; Length 202;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 22 LAASLSA 28
Db 12 LAASLSA 18
```

```
Search completed: March 20, 2005, 10:19:23
Job time : 28 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 20, 2005, 10:18:25 ; Search time 73 Seconds
(without alignments)
1122.934 Million cell updates/sec

Title: US-10-080-522-1

Perfect score: 248

Sequence: 1 MQTCPLAPGCHVSQALGTL.....PLGALELLSPQLPPYAADP 248

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 1401741 seqs, 330541175 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1401741

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications AA.*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	248	100.0	248	9 US-09-799-777-32	Sequence 32, Appl
2	248	100.0	248	9 US-09-997-165-4	Sequence 4, Appl
3	248	100.0	248	14 US-10-080-522-1	Sequence 1, Appl
4	9	3.6	1709	14 US-10-156-761-11180	Sequence 11180, A
5	8	3.2	75	16 US-10-437-963-110606	Sequence 110606,
6	8	3.2	123	16 US-10-437-963-116750	Sequence 116750,
7	8	3.2	125	15 US-10-424-599-182088	Sequence 182088,
8	8	3.2	294	15 US-10-282-122A-48538	Sequence 48538, A
9	8	3.2	384	15 US-10-282-122A-76494	Sequence 76494, A
10	8	3.2	545	14 US-10-156-761-11660	Sequence 11660, A
11	8	3.2	691	15 US-10-369-493-10173	Sequence 10173, A
12	7	2.8	19	15 US-10-468-372-9	Sequence 9, Appl
13	7	2.8	36	11 US-09-833-245-858	Sequence 858, App

Sequence 29093, A
Sequence 50684, A
Sequence 2378, Ap
Sequence 195864,
Sequence 145761,
Sequence 134138,
Sequence 342, App
Sequence 170987,
Sequence 9429, Ap
Sequence 3112, Ap
Sequence 2594, Ap
Sequence 176, App
Sequence 252, App
Sequence 222, App
Sequence 222, App
Sequence 232, App
Sequence 4, Appl
Sequence 160, App
Sequence 160, App
Sequence 39979, A
Sequence 165259,
Sequence 2088, Ap
Sequence 2, Appl
Sequence 1399, Ap
Sequence 2323, Ap
Sequence 184172,
Sequence 155, App
Sequence 155, App
Sequence 175211,
Sequence 175849,
Sequence 132222,

US-10-029-386-29093
US-10-767-701-50684
US-09-864-408A-2978
US-10-437-963-195864
US-10-424-599-145761
US-10-437-963-124138
US-10-195-730-342
US-10-799-747-342
US-10-437-963-170987
US-10-156-761-9429
US-10-104-047-3112
US-10-108-260A-2594
US-09-881-752A-176
US-09-731-872-252
US-09-981-876-222
US-09-148-545-222
US-09-876-997-252
US-10-468-372-4
US-09-981-876-160
US-09-148-545-160
US-10-767-701-39979
US-10-437-963-184172
US-10-097-065-155
US-10-372-876-155
US-10-437-963-175211
US-10-424-599-175849
US-10-437-963-132222

ALIGNMENTS

RESULT 1

US-09-799-777-32
; Sequence 32, Application US/09799777
; Patent No. US20020091244A1
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; Hillman, Jennifer L.
; Corley, Neil C.
; Guegler, Karl J.
; Baugh, Mariah
; Sather, Susan
; Shah, Purvi
; TITLE OF INVENTION: HUMAN SIGNAL PEPTIDE-CONTAINING PROTEINS
; NUMBER OF SEQUENCES: 154
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/799,777
; FILING DATE: 06-Mar-2001
; PRIORITY DATA:
; CLASSIFICATION: <Unknown>
; APPLICATION NUMBER: US/09/002,485
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: BILLINGS, LUCY J.
; REGISTRATION NUMBER: 36,749

us-10-080-522-1-oli.rapb

Sun Mar 20 14:17:41 2005

```

; REFERENCE/DOCKET NUMBER: PF-0459 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 248 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 32 :
US-09-799-777-32

Query Match      100.0%; Score 248; DB 9; Length 248;
Best Local Similarity 100.0%; Pred. No. 1.8e-230; Indels 0; Gaps 0;
Matches 248; Conservative 0; Mismatches 0;

QY 1 MOTCPFLAFPGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSCNISNA 60
DB 1 MOTCPFLAFPGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSCNISNA 60

QY 61 FSHVNIKLRAHQESAI FNEVAPGYFSRDGWLQVGGVAQLVIKGA RDSHAGLYMHLV 120
DB 61 FSHVNIKLRAHQESAI FNEVAPGYFSRDGWLQVGGVAQLVIKGA RDSHAGLYMHLV 120

QY 121 GHORNNRQVTLVSGAEPOSADPTGFWPVPVAVTVFILLVALVMFAYRCRCSQORREK 180
DB 121 GHORNNRQVTLVSGAEPOSADPTGFWPVPVAVTVFILLVALVMFAYRCRCSQORREK 180

QY 181 KFFLLPQMKVAALRAGAQQGLSRASAEIWTDPSEPTPRPLALVPKPSPLGALLELLSQP 240
DB 181 KFFLLPQMKVAALRAGAQQGLSRASAEIWTDPSEPTPRPLALVPKPSPLGALLELLSQP 240

QY 241 LFPYAADP 248
DB 241 LFPYAADP 248

RESULT 2
US-09-997-165-4
; Sequence 4, Application US/09997165
; Patent No. US20020141999A1
; GENERAL INFORMATION:
; APPLICANT: Lyman, Stewart D.
; APPLICANT: Fanslow, William C.
; TITLE OF INVENTION: LIGAND FOR CD7 AND METHODS OF USE THEREOF
; FILE REFERENCE: 2913-US
; CURRENT APPLICATION NUMBER: US/09/997,165
; CURRENT FILING DATE: 2001-11-27
; PRIOR APPLICATION NUMBER: PCT/US00/14612
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/136,450
; PRIOR FILING DATE: 1999-05-28
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 248
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-997-165-4

Query Match      100.0%; Score 248; DB 9; Length 248;
Best Local Similarity 100.0%; Pred. No. 1.8e-230; Indels 0; Gaps 0;
Matches 248; Conservative 0; Mismatches 0;

QY 1 MOTCPFLAFPGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSCNISNA 60
DB 1 MOTCPFLAFPGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSCNISNA 60

QY 61 FSHVNIKLRAHQESAI FNEVAPGYFSRDGWLQVGGVAQLVIKGA RDSHAGLYMHLV 120
DB 61 FSHVNIKLRAHQESAI FNEVAPGYFSRDGWLQVGGVAQLVIKGA RDSHAGLYMHLV 120

; REFERENCE/DOCKET NUMBER: PF-0459 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 248 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 32 :
US-10-080-522-1

Query Match      100.0%; Score 248; DB 14; Length 248;
Best Local Similarity 100.0%; Pred. No. 1.8e-230; Indels 0; Gaps 0;
Matches 248; Conservative 0; Mismatches 0;

QY 1 MOTCPFLAFPGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSCNISNA 60
DB 1 MOTCPFLAFPGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSCNISNA 60

QY 61 FSHVNIKLRAHQESAI FNEVAPGYFSRDGWLQVGGVAQLVIKGA RDSHAGLYMHLV 120
DB 61 FSHVNIKLRAHQESAI FNEVAPGYFSRDGWLQVGGVAQLVIKGA RDSHAGLYMHLV 120

; SEQUENCE 1, Application US/10080522
; Publication No. US20030096326A1
; GENERAL INFORMATION:
; APPLICANT: KAUFMAN, RUSSEL E.
; APPLICANT: SLENTZ-KESLER, KIMBERLY
; TITLE OF INVENTION: GENE PRODUCT OVER EXPRESSED IN CANCER
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHUYE P.C.
; STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/080,522
; FILING DATE: 25-Feb-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/539,774
; FILING DATE: 31-MAR-2000
; APPLICATION NUMBER: US 09/210,474
; FILING DATE: 14-DEC-1998
; APPLICATION NUMBER: US 08/755,559
; FILING DATE: 22-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: WILSON, MARY J.
; REGISTRATION NUMBER: 32,955
; REFERENCE/DOCKET NUMBER: 1579-645
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 248 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-080-522-1

```

Db 61 FSHVNIKLRAGHGESAIENFVAPGYFSRDRGWLQVGGVQAQLVVIKARDSHAGLYMHV 120
QY 121 GQRNNRQVTLVSVGAEPQSDPTGFWPVPVAVTAVFILLVALVMFAWYRCRCQORREK 180
Db 121 GQRNNRQVTLVSVGAEPQSDPTGFWPVPVAVTAVFILLVALVMFAWYRCRCQORREK 180
QY 181 KFFLLEPQMKVAALRAGAQQGLSRASAEIWTDPSEPTRPLALVPKPSPLGALELLSPQ 240
Db 181 KFFLLEPQMKVAALRAGAQQGLSRASAEIWTDPSEPTRPLALVPKPSPLGALELLSPQ 240
QY 241 LPFYAADP 248
Db 241 LPFYAADP 248

RESULT 4
US-10-156-761-11180
; Sequence 11180, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 11180
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-11180

Query Match 3.6%; Score 9; DB 14; Length 1709;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 229 PLGALELLS 237
Db 1063 PLGALELLS 1071

RESULT 5
US-10-437-963-110606
; Sequence 110606, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 110606
; LENGTH: 75
; TYPE: PRT

; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_14657C.1.pap
US-10-437-963-110606

Query Match 3.2%; Score 8; DB 16; Length 75;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 153 VTAVFILL 160
Db 58 VTAVFILL 65

RESULT 6
US-10-437-963-116750
; Sequence 116750, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 116750
; LENGTH: 123
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(123)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_20220C.1.pap
US-10-437-963-116750

Query Match 3.2%; Score 8; DB 16; Length 123;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 230 LGALELLS 237
Db 21 LGALELLS 28

RESULT 7
US-10-424-599-182088
; Sequence 182088, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 182088
; LENGTH: 125
; TYPE: PRT
; ORGANISM: Glycine max

APPLICANT: HATTORI, MASAIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 11660
LENGTH: 545
TYPE: PRT
ORGANISM: Streptomyces avermitilis
US-10-156-761-11660

Query Match 3.2%; Score 8; DB 14; Length 545;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 SPLGALEL 235
DB 120 SPLGALEL 127

RESULT 11
US-10-369-493-10173
Sequence 10173, Application US/10369493
Publication No. US20030233675A1
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xiaofeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 10173
LENGTH: 691
TYPE: PRT
ORGANISM: magnetite-containing magnetic coccus
US-10-369-493-10173

Query Match 3.2%; Score 8; DB 15; Length 691;
Best Local Similarity 100.0%; Pred. No. 93;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LGTLLFLA 23
DB 7 LGTLLFLA 14

RESULT 12
US-10-468-372-9
Sequence 9, Application US/10468372
Publication No. US20040081984A1
GENERAL INFORMATION:
APPLICANT: The University of Virginia Patent Foundation
APPLICANT: Laurie, Gordon W
APPLICANT: Kumar, Rajesh W
APPLICANT: Sanghi, Sandhya
APPLICANT: Lumsden, Angela
TITLE OF INVENTION: Ocular Tear Growth Factor-Like Protein
FILE REFERENCE: 00662-02
CURRENT APPLICATION NUMBER: US/10/468,372
CURRENT FILING DATE: 2003-08-19
PRIOR APPLICATION NUMBER: US 60/269,900

PRIOR FILING DATE: 2001-02-20
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn version 3.1
SEQ ID NO 9
LENGTH: 19
TYPE: PRT
ORGANISM: Homo sapiens
US-10-468-372-9

Query Match 2.8%; Score 7; DB 15; Length 19;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 TLLFLAA 24
DB 5 TLLFLAA 11

RESULT 13
US-09-833-245-858
Sequence 858, Application US/09833245
Publication No. US20040010134A1
GENERAL INFORMATION:
APPLICANT: Human Genome Sciences, Inc.
TITLE OF INVENTION: Albumin Fusion Proteins
FILE REFERENCE: PF546PCT
CURRENT APPLICATION NUMBER: US/09/833,245
CURRENT FILING DATE: 2001-04-12
PRIOR APPLICATION NUMBER: 60/229, 358
PRIOR FILING DATE: 2000-04-12
PRIOR APPLICATION NUMBER: 60/256, 931
PRIOR FILING DATE: 2000-12-21
PRIOR APPLICATION NUMBER: 60/199, 384
PRIOR FILING DATE: 2000-04-25
NUMBER OF SEQ ID NOS: 2267
SOFTWARE: PatentIn ver. 2.1
SEQ ID NO 858
LENGTH: 36
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (17)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-833-245-858

Query Match 2.8%; Score 7; DB 11; Length 36;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LLFLAAS 25
DB 18 LLFLAAS 24

RESULT 14
US-10-029-386-29093
Sequence 29093, Application US/10029386
Publication No. US20030194704A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharron G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: AEOMICA-X-2
CURRENT APPLICATION NUMBER: US/10/029,386
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 34288
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 29093
LENGTH: 56
TYPE: PRT

Sun Mar 20 14:17:41 2005

```

; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR19.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 5.8
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.6
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 6.8
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.8
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 5.3
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.3
US-10-029-386-29093

```

```

Query Match      2.8%  Score 7; DB 14; Length 56;
Best Local Similarity 100.0%; Pred. No. 91;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 133 VSGAEPQ 139
    |||||
DB 35 VSGAEPQ 41

```

```

RESULT 15
US-10-767-701-50684
; Sequence 50684, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 50684
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3480-024-PI-K1-H1.pep
US-10-767-701-50684

```

```

Query Match      2.8%  Score 7; DB 16; Length 58;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 181 KFFLLEP 187
    |||||
DB 15 KFFLLEP 21

```

Search completed: March 20, 2005, 10:25:26
Job time : 76 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 20, 2005, 10:13:31 ; Search time 25 Seconds
(without alignments)
954.470 Million cell updates/sec

Title: US-10-080-522-1
Perfect score: 248
Sequence: 1 MQTCPLAPGHVSQALGTL.....PLGALELLSQPLFPYAADP 248

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283416 seqs, 96216763 residues
Word size : 0
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	IP	Description
1	8	3.2	108	2	F72549	hypothetical prote
2	8	3.2	384	2	E71331	probable cell divi
3	8	3.2	10797	2	T30192	probable peptide s
4	7	2.8	51	2	C82707	hypothetical prote
5	7	2.8	74	2	T16336	hypothetical prote
6	7	2.8	96	2	T22979	hypothetical prote
7	7	2.8	131	2	AE0126	conserved hypothet
8	7	2.8	131	4	S63446	hypothetical prote
9	7	2.8	134	2	D64558	hypothetical prote
10	7	2.8	136	2	S37924	hypothetical prote
11	7	2.8	143	2	A97133	uncharacterized ph
12	7	2.8	156	2	F82575	DNA methylase Xf22
13	7	2.8	162	2	S70593	NADH2 dehydrogenas
14	7	2.8	180	1	JQ0986	lipid body-associat
15	7	2.8	181	2	AE0435	conserved hypothet
16	7	2.8	181	2	B82134	transcriptional regu
17	7	2.8	195	2	A99346	hypothetical prote
18	7	2.8	207	2	B44994	eggshell protein 2
19	7	2.8	212	2	A44994	hypothetical prote
20	7	2.8	225	2	T25075	hypothetical prote
21	7	2.8	229	2	F70568	hypothetical prote
22	7	2.8	234	2	C81149	hypothetical prote
23	7	2.8	237	2	G84678	probable RING zinc
24	7	2.8	243	2	B82574	site-specific DNA-
25	7	2.8	246	2	G83959	tRNA methyltransfe
26	7	2.8	248	2	H70511	probable prcA prot
27	7	2.8	259	2	F82087	deoxyribose-phosph
28	7	2.8	265	2	E87074	proteasome [alpha]
29	7	2.8	265	2	S72865	hypothetical prote

30	7	2.8	272	2	A69961	geranyltranstransf
31	7	2.8	276	2	E96951	conserved membrane
32	7	2.8	284	1	E64938	probable aldehyde
33	7	2.8	284	2	F85788	probable an aldehy
34	7	2.8	284	2	B90940	probable an aldehy
35	7	2.8	288	1	S70955	otnf protein - Vib
36	7	2.8	288	2	T44314	hypothetical prote
37	7	2.8	288	2	AE2947	ABC transporter, m
38	7	2.8	290	2	G75470	conserved hypothet
39	7	2.8	290	2	AE1126	hypothetical prote
40	7	2.8	299	2	B70737	hypothetical prote
41	7	2.8	304	2	B92224	probable permease
42	7	2.8	304	2	AE3062	hypothetical prote
43	7	2.8	307	2	G82385	hypothetical prote
44	7	2.8	311	2	E98335	membrane protein n
45	7	2.8	312	2	AD2797	malonate transport

ALIGNMENTS

RESULT 1

F72549

hypothetical protein APE1682 - Aeropyrum pernix (strain K1)

C:Species: Aeropyrum pernix

C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004

C:Accession: F72549

R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah

awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K

DNA Res. 6, 83-101, 1999

A>Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropy

A:Reference number: A72450; MUID:99310339; PMID:10382966

A:Accession: F72549

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-108 <KAW>

A:Cross-references: UNIPROT:Q9YBB5; DBJ:AP000062; NID:G5105244; PIDN:BAA80683.1; PID:G5

A:Experimental source: strain K1

C:Genetics:

A:Gene: APE1682

C:Superfamily: Aeropyrum pernix hypothetical protein APE1682

Query Match 3.2%; Score 8; DB 2; Length 108;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 217 TPRPLALV 224

Db 19 TPRPLALV 26

RESULT 2

E71331

probable cell division protein (ftsW) - syphilis spirochete

C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)

C>Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 09-Jul-2004

C:Accession: E71331

R:Fraser, C.M.; Norris, S.J.; Weinstein, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin

ston, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McD

they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.

Science 281, 375-388, 1998

A>Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.

A:Reference number: A71250; MUID:98332770; PMID:9665876

A:Accession: E71331

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-384 <COL>

A:Cross-references: UNIPROT:O83402; GB:AE001217; GB:AE000520; NID:G3322656; PIDN:AAC6537

A:Experimental source: strain Nichols

C:Genetics:

A:Gene: TP0387

C:Superfamily: rod shape-determining protein

Query Match 3.2%; Score 8; DB 2; Length 384;
Best Local Similarity 100.0%; Pred. No. 7.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 156 VFILVAL 163
Db 19 VFILVAL 26

RESULT 3
T30192
probable peptide synthetase - Aureobasidium pullulans
C:Species: Aureobasidium pullulans
C:Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 09-Jul-2004
C:Accession: T30192
R:Peery, R.B.; Thornwell, S.J.; Tobin, M.B.; Skatrud, P.L.
submitted to the EMBL Data Library, January 1997
A:Description: Discovery of an MDR-like gene adjacent to a peptide synthetase in Aureoba
A:Reference number: Z20767
A:Accession: T30192
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-10797 <P>
A:Cross-references: UNIPROT:Q94116; EMBL:U85909; NID:G4099310; PID:G4099313; PIDN:AD005
C:Genetics:
A:Introns: 2078/3; 4142/3; 6075/3; 7963/3; 7985/2
C:Keywords: carrier protein
F:1618-1688/Domain: acyl carrier protein homology <ACP1>
F:1682-3752/Domain: acyl carrier protein homology <ACP2>
F:5615-5685/Domain: acyl carrier protein homology <ACP3>
F:7503-7573/Domain: acyl carrier protein homology <ACP4>
F:9683-9752/Domain: acyl carrier protein homology <ACP5>

Query Match 3.2%; Score 8; DB 2; Length 10797;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 230 LGALELLS 237
Db 598 LGALELLS 605

RESULT 4
C82707
hypothetical protein XF1235 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: C82707
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717; PMID:10910347
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: C82707
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-51 <SIM>
A:Cross-references: UNIPROT:Q9PDZ3; GB:AE003957; GB:AE003849; NID:G9106207; PIDN:AAF8404
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briñones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Prohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krüger, J.E.; Kuramae, E.E.; Laigh
Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tshunako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A:Reference number: A59328
A:Contents: annotation

C:Genetics:
A:Gene: XF1235

Query Match 2.8%; Score 7; DB 2; Length 51;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 191 VAALRAG 197
Db 12 VAALRAG 18

RESULT 5
T16336
hypothetical protein F41G3.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C:Accession: T16336
R:Menezes, S.
submitted to the EMBL Data Library, November 1995
A:Description: The sequence of C. elegans cosmid F41G3.
A:Reference number: Z18496
A:Accession: T16336
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-74 <MEN>
A:Cross-references: UNIPROT:Q20295; EMBL:U39999; NID:G1055133; PID:G1055142; PIDN:AAA811
C:Genetics:
A:Gene: CESP:F41G3.1
A:Introns: 22/1

Query Match 2.8%; Score 7; DB 2; Length 74;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 201 GLSRASA 207
Db 61 GLSRASA 67

RESULT 6
T22979
hypothetical protein F59B10.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T22979
R:Lloyd, C.
submitted to the EMBL Data Library, March 1995
A:Reference number: Z19646
A:Accession: T22979
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-96 <WIL>
A:Cross-references: UNIPROT:Q09953; EMBL:Z48716; PIDN:CAA88599.1; GSPDB:GN000020; CESP:F
A:Experimental source: clone F59B10
C:Genetics:
A:Gene: CESP:F59B10.5
A:Map position: 2
A:Introns: 33/3; 73/2

Query Match 2.8%; Score 7; DB 2; Length 96;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 157 FILLVAL 163
Db 4 FILLVAL 10

RESULT 7
AE0126
conserved hypothetical protein YP01030 [imported] - Yersinia pestis (strain CO92)
C:Species: Yersinia pestis

C>Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
 C/Accession: AE0126
 R/Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall, Nature 413, 523-527, 2001
 A>Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.
 A/Reference number: AB0001; MUID:21470413; PMID:11586360
 A/Accession: AE0126
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-131 <CUR>
 A/Cross-references: UNIPROT:Q8ZH79; GB:AL590842; PIDN:CAC98872.1; PID:g15979097; GSPDB:C
 C/Genetics:
 A:Gene: YPO1030

Query Match 2.8%; Score 7; DB 2; Length 131;
 Best Local Similarity 100.0%; Pred. No. 32;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 ALGTLFLF 21
 Db 77 ALGTLFLF 83

RESULT 8
 S63446
 hypothetical protein pEOA461 - yeast (*Saccharomyces cerevisiae*)
 C/Species: *Saccharomyces cerevisiae*
 C/Date: 16-May-1996 #sequence_revision 05-Sep-1996 #text_change 09-Jul-2004
 C/Accession: S71984; S63446
 R/Casamayor, A.; Khalid, H.; Baicells, L.; Aldea, M.; Casas, C.; Herrero, E.; Arino, J. Yeast 12, 1013-1020, 1996
 A>Title: Sequence analysis of a 13.4 kbp fragment from the left arm of chromosome XV revealing frames.
 A/Reference number: S71978; MUID:97051588; PMID:8896265
 A/Accession: S71984
 A/Status: nucleic acid sequence not shown; translation not shown; conceptual translation
 A/Molecule type: DNA
 A/Residues: 1-131 <CAS>
 A/Cross-references: UNIPROT:Q02812; EMBL:U01293; NID:g1209710; PIDN:AAC49468.1; PID:g120
 A/Experimental source: strain F1679
 C/Note: The nucleotide sequence was submitted to the EMBL Data Library, November 1995
 C/Comment: There is no evidence that this sequence is expressed.
 C/Genetics:
 A:Map position: 15L
 C/Keywords: pseudogene

Query Match 2.8%; Score 7; DB 4; Length 131;
 Best Local Similarity 100.0%; Pred. No. 32;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 157 FILLVAL 163
 Db 97 FILLVAL 103

RESULT 9
 D64558
 hypothetical protein HP0308 - *Helicobacter pylori* (strain 26695)
 C/Species: *Helicobacter pylori*
 C/Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 09-Jul-2004
 C/Accession: D64558
 R/Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodet, A.; McKenne son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L. Nature 388, 539-547, 1997
 A/Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C. A>Title: The complete genome sequence of the gastric pathogen *Helicobacter pylori*.
 A/Reference number: A64520; MUID:97394467; PMID:9252185
 A/Accession: D64558
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA

A/Residues: 1-134 <TOM>
 A/Cross-references: UNIPROT:O25078; GB:AE000549; GB:AE000511; NID:g2313403; PIDN:AAD0738

Query Match 2.8%; Score 7; DB 2; Length 134;
 Best Local Similarity 100.0%; Pred. No. 32;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LGTLLFLF 22
 Db 97 LGTLLFLF 103

RESULT 10
 S37924
 hypothetical protein YKL097c - yeast (*Saccharomyces cerevisiae*)
 A/Alternate names: hypothetical protein YKL445
 C/Species: *Saccharomyces cerevisiae*
 C/Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 09-Jul-2004
 C/Accession: S37924; S39088
 R/Cheret, G.; Fukuhara, H.; Bolotin-Fukuhara, M.; Daignan-Fornier, B.; Pallier, C.; Puzo submitted to the Protein Sequence Database, March 1994
 A/Reference number: S37920
 A/Accession: S37924
 A/Molecule type: DNA
 A/Residues: 1-136 <CHR>
 A/Cross-references: UNIPROT:P34245; EMBL:Z28097; NID:g486160; PIDN:CAA81936.1; PID:g4861
 A/Experimental source: strain S288C
 R/Pallier, C.; Valens, M.; Puzos, V.; Fukuhara, H.; Cheret, G.; Sor, F.; Bolotin-Fukuhar Yeast 9, 1149-1155, 1993
 A>Title: DNA sequence analysis of a 17 kb fragment of yeast chromosome XI physically loc protein kinases.
 A/Reference number: S39084; MUID:94078677; PMID:8256524
 A/Accession: S39088
 A/Status: translation not shown
 A/Molecule type: DNA
 A/Residues: 1-136 <PAL>
 A/Cross-references: EMBL:X71133; NID:g431205; PIDN:CAA50460.1; PID:g431219
 A/Experimental source: strain S288C
 C/Genetics:
 A:Cross-references: SGD:S0001580
 A:Map position: 11L
 C/Superfamily: *Saccharomyces* hypothetical protein YKL097c

Query Match 2.8%; Score 7; DB 2; Length 136;
 Best Local Similarity 100.0%; Pred. No. 33;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 147 WPVPAVV 153
 Db 17 WPVPAVV 23

RESULT 11
 A97133
 uncharacterized phage related protein [imported] - *Clostridium acetobutylicum*
 C/Species: *Clostridium acetobutylicum*
 C/Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 09-Jul-2004
 C/Accession: A97133
 R/Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R. J. Bacteriol. 183, 4823-4838, 2001
 A>Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium *Clostridium acetobutylicum* ATCC824
 A/Reference number: A96900; MUID:21359325; PMID:21359325
 A/Accession: A97133
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-143 <CUR>
 A/Cross-references: UNIPROT:Q97HW8; GB:AB001437; PIDN:AAK79852.1; PID:g15024868; GSPDB:G
 A/Experimental source: *Clostridium acetobutylicum* ATCC824
 C/Genetics:
 A:Gene: CAC1888

Query Match 2.8%; Score 7; DB 2; Length 143;

Query Match 2.8%; Score 7; DB 2; Length 156;
 Best Local Similarity 100.0%; Pred. No. 34;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 193 ALRGAQ 199
 |||||
 Db 31 ALRGAQ 37

RESULT 12
 DNA methylase XP2297 [imported] - Xylella fastidiosa (strain 9a5c)
 C:Species: Xylella fastidiosa
 C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
 C:Accession: F82575
 R:Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing
 Nature 406, 151-157, 2000
 A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
 A:Reference number: A82515; MUID:20365717; PMID:10910347
 A>Note: for a complete list of authors see reference number A59328 below
 A:Accession: F82575
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-156 <S1M>
 A:Cross-references: UNIPROT:Q9PB48; GB:AE004041; GB:AE003849; NID:g9107453; PIDN:AAF8509
 A:Experimental source: strain 9a5c
 R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
 Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrier, H
 as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
 submitted to GenBank, June 2000
 A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
 J.D.; Junqueira, M.L.; Kemper, E.B.; Kicajina, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig
 chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.B.; Marques, M.V.; Martins, E
 A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
 F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
 Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
 A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
 M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
 A:Reference number: A59328
 A:Contents: annotation
 C:Genetics:
 A:Gene: XP2297

Query Match 2.8%; Score 7; DB 2; Length 156;
 Best Local Similarity 100.0%; Pred. No. 37;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 191 VAALRAG 197
 |||||
 Db 136 VAALRAG 142

RESULT 13
 NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 6 - starfish (Asterina pectinifera)
 C:Species: mitochondrion Asterina pectinifera
 C>Date: 14-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
 C:Accession: S70593
 R:Asakawa, S.; Himeno, H.; Miura, K.; Watanabe, K.
 Genetics 140, 1047-1060, 1995
 A:Title: Nucleotide sequence and gene organization of the starfish Asterina pectinifera
 A:Reference number: S70593; MUID:95402698; PMID:7672576
 A:Accession: S70593
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-162 <ASA>
 A:Cross-references: UNIPROT:Q33817; EMBL:D16387; NID:9517153; PIDN:BAR03876.1; PID:g1060
 A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1993
 C:Genetics:
 A:Genome: mitochondrion
 A:Genetic code: SGC8
 C:Superfamily: NADH dehydrogenase (ubiquinone) chain 6
 C:Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Query Match 2.8%; Score 7; DB 2; Length 162;
 Best Local Similarity 100.0%; Pred. No. 38;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 158 ILLVALV 164
 |||||
 Db 139 ILLVALV 145

RESULT 14
 JQ0986
 Lipid body-associated membrane protein - carrot
 C:Species: Daucus carota (carrot)
 C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
 C:Accession: JQ0986
 R:Hatzopoulos, P.; Franz, G.; Choy, L.; Sung, R.Z.
 Plant Cell 2, 457-467, 1990
 A:Title: Interaction of nuclear factors with upstream sequences of a lipid body membrane
 A:Reference number: JQ0986; MUID:93044488; PMID:2152171
 A:Accession: JQ0986
 A:Molecule type: DNA
 A:Residues: 1-180 <HAT>
 A:Cross-references: UNIPROT:Q43123; GB:S47635; NID:g259452; PIDN:AAB24078.1; PID:g259453
 A:Experimental source: var. juwarot
 C:Genetics:
 A:Gene: DC59
 C:Superfamily: oleosin
 C:Keywords: membrane protein

Query Match 2.8%; Score 7; DB 1; Length 180;
 Best Local Similarity 100.0%; Pred. No. 42;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 GTLLFLA 23
 |||||
 Db 50 GTLLFLA 56

RESULT 15
 AE0435
 conserved hypothetical protein YP03580 [imported] - Yersinia pestis (strain CO92)
 C:Species: Yersinia pestis
 C>Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
 C:Accession: AE0435
 R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B
 geno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
 il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
 Nature 413, 523-527, 2001
 A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
 A:Reference number: AB0001; MUID:21470413; PMID:11586360
 A:Accession: AE0435
 A:Molecule type: DNA
 A>Status: preliminary
 A:Residues: 1-181 <KUR>
 A:Cross-references: UNIPROT:Q8ZB45; GB:AL590842; PIDN:CAC92809.1; PID:g15981502; GSPDB:G
 C:Genetics:
 A:Gene: YP03580
 C:Superfamily: conserved hypothetical protein NMB0355

Query Match 2.8%; Score 7; DB 2; Length 181;
 Best Local Similarity 100.0%; Pred. No. 42;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 LAASLSA 28
 |||||
 Db 17 LAASLSA 23

Search completed: March 20, 2005, 10:18:52
 Job time : 29 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 20, 2005, 09:50:19 ; Search time 92 Seconds
(without alignments)
1380.387 Million cell updates/sec

Title: US-10-080-522-1

Perfect score: 248

Sequence: 1 MQTCPLAPGHVSQALGTL.....PLGALELLSQPLFPYAADP 248

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 1612378 seqs, 512079187 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	248	100.0	248	1	SCTM HUMAN
2	9	3.6	1709	2	Q8WV66 homo sapien
3	8	3.2	108	2	Q93H59 streptomyc
4	8	3.2	108	2	Q9YBB5 aeropyrum p
5	8	3.2	124	2	Q8R155 mus musculu
6	8	3.2	159	2	Q6Y402 ophiura lut
7	8	3.2	163	2	Q8X1D5 uestilago ma
8	8	3.2	201	2	Q6AYS0 Q8X80 rattus norv
9	8	3.2	221	2	Q7NUR2 chromobacte
10	8	3.2	288	2	Q659M6 turnip mosa
11	8	3.2	304	2	Q64WT9 bacteroides
12	8	3.2	316	2	Q6P2X0 xenopus tro
13	8	3.2	332	2	Q7Y516 xanthomonas
14	8	3.2	384	2	Q83402 treponema p
15	8	3.2	521	2	Q897Y6 clostridium
16	8	3.2	545	2	Q82FX6 streptomyc
17	8	3.2	760	2	Q9N612 leishmania
18	8	3.2	2204	2	Q6Q356 newcastle d
19	8	3.2	2204	2	Q6RCH6 newcastle d
20	7	2.8	39	2	Q94116 aureobasidi
21	7	2.8	51	2	Q9ERK6 mus musculu
22	7	2.8	56	2	Q9PD23 xyella fas
23	7	2.8	59	2	Q7Y191 oryza sativ
24	7	2.8	59	2	Q745H6 mycobacteri
25	7	2.8	68	1	Q6MMJ7 opisthacant
26	7	2.8	70	2	Q8UX21 hepatitis c
27	7	2.8	70	2	Q8UX22 hepatitis c
28	7	2.8	70	2	Q8UX23 hepatitis c
29	7	2.8	70	2	Q8UX24 hepatitis c
30	7	2.8	70	2	Q8UX25 hepatitis c
31	7	2.8	70	2	Q8UX27 hepatitis c
32	7	2.8	70	2	Q8UX28 hepatitis c

32	7	2.8	70	2	Q8UX29	Q8ux29 hepatitis c
33	7	2.8	70	2	Q8UX30	Q8ux30 hepatitis c
34	7	2.8	70	2	Q8UX31	Q8ux31 hepatitis c
35	7	2.8	70	2	Q8UX32	Q8ux32 hepatitis c
36	7	2.8	70	2	Q8UX33	Q8ux33 hepatitis c
37	7	2.8	70	2	Q8UX34	Q8ux34 hepatitis c
38	7	2.8	70	2	Q8UX35	Q8ux35 hepatitis c
39	7	2.8	70	2	Q8UX36	Q8ux36 hepatitis c
40	7	2.8	70	2	Q8UX37	Q8ux37 hepatitis c
41	7	2.8	70	2	Q8UX38	Q8ux38 hepatitis c
42	7	2.8	70	2	Q8UX39	Q8ux39 hepatitis c
43	7	2.8	70	2	Q8UX40	Q8ux40 hepatitis c
44	7	2.8	70	2	Q8UX41	Q8ux41 hepatitis c
45	7	2.8	70	2	Q8UX42	Q8ux42 hepatitis c
					Q8UX43	Q8ux43 hepatitis c

ALIGNMENTS

RESULT 1
SCTM HUMAN STANDARD; PRT; 248 AA.
AC Q8WV66; Q00466;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Secreted and transmembrane protein 1 precursor (Protein K12).
GN Name=SECTM1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haiech F.,
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Fahey J., Helton E., Kettelman M., Madan A.C., Rodrigues S., Sanchez A.,
Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [3]
RP INTERACTION WITH CD7.
RX MEDLINE=20119303; PubMed=10652336; DOI=10.1074/jbc.275.5.3431;
RA Lyman S.D., Escobar S., Rousseau A.-M., Armstrong A., Fanslow W.C.;
RT "Identification of CD7 as a cognate of the human K12 (SECTM1)
protein";
RL J. Biol. Chem. 275:3431-3437 (2000).
CC -!- SUBUNIT: Interacts with CD7.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Probable). Also

CC found as secreted.
 CC -1- TISSUE SPECIFICITY: Detected at the highest levels in peripheral
 CC blood leukocytes and breast cancer cell lines. Found in leukocytes
 CC of the myeloid lineage, with the strongest expression observed in
 CC granulocytes and no detectable expression in lymphocytes.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U77643; AAC52044.1; --
 CC EMBL; BC017716; AAH17716.1; --
 CC Genew; HGNC:10707; SECTW1.
 CC H-InvDB; HIX0014254; --
 CC MIM; 602602; --
 CC InterPro; IPR007110; Ig-like.
 CC Signal; Transmembrane.
 CC SIGNAL 1 28 Potential.
 CC CHAIN 29 248 Secreted and transmembrane protein 1.
 CC DOMAIN 29 145 Extracellular (Potential).
 CC TRANSMEM 146 166 Potential.
 CC DOMAIN 167 248 Cytoplasmic (Potential).
 CC DISULFID 38 55 Potential.
 CC CARBOHYD 56 56 N-linked (GlcNAc...) (Potential).
 CC CONFLICT 191 191 V -> F (in Ref. 2).
 CC SEQUENCE 248 AA; 27039 MW; 21E3066B67920487 CRC64;
 CC
 CC Query Match 100.0%; Score 248; DB 1; Length 248;
 CC Best Local Similarity 100.0%; Pred. No. 3.9e-240;
 CC Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC QY 1 MQTCPLAFCHVSQALGTLLELAASLAQNEGWDSPTCTGVVSVSWGNTWVSCNLSNA 60
 CC Db 1 MQTCPLAFPHGVHVSQALGTLLELAASLAQNEGWDSPTCTGVVSVSWGNTWVSCNLSNA 60
 CC
 CC QY 61 FSHVNIKLRHAGQESAI FNEVAFGYFSDGWLQVGGVAQLVKGARDSHAGLYMWHLV 120
 CC Db 61 FSHVNIKLRHAGQESAI FNEVAFGYFSDGWLQVGGVAQLVKGARDSHAGLYMWHLV 120
 CC
 CC QY 121 GHORNNRQVTVLGVSGAPDPTGFWVPVAVTVFILLVAVMPFAWTRCSCQORREK 180
 CC Db 121 GHORNNRQVTVLGVSGAPDPTGFWVPVAVTVFILLVAVMPFAWTRCSCQORREK 180
 CC
 CC QY 181 KFFLEPQMKVAALRAGAQOGLSRASAEIWTPTDSEPTPRPLALVFKPSPGLALELLSPQ 240
 CC Db 181 KFFLEPQMKVAALRAGAQOGLSRASAEIWTPTDSEPTPRPLALVFKPSPGLALELLSPQ 240
 CC
 CC QY 241 LFPYAADP 248
 CC Db 241 LFPYAADP 248
 CC
 CC RESULT 2
 CC ID Q93H59 PRELIMINARY; PRT; 1709 AA.
 CC AC Q93H59
 CC DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 CC DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 CC DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 CC DE Non-ribosomal peptide synthetase (Putative non-ribosomal peptide
 CC synthetase).
 CC GN OrderedLocusNames=SAV3643;
 CC OS Streptomyces avermitilis.
 CC OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 CC OC Streptomycineae; Streptomycetaceae; Streptomyces.
 CC OX NCBI_TaxID=33903;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN=MA-4680;

RX MEDLINE-21477403; PubMed=11572948; DOI=10.1073/pnas.211433198;
 RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
 RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
 RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
 RT "Genome sequence of an industrial microorganism Streptomyces
 RT avermitilis; deducing the ability of producing secondary
 RT metabolites.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
 RL [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MA-4680;
 RX MEDLINE-22608306; PubMed=12692562;
 RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
 RA Sakaki Y., Hattori M., Omura S.;
 RT "Complete genome sequence and comparative analysis of the industrial
 RT microorganism Streptomyces avermitilis.";
 RL Nat. Biotechnol. 21:526-531(2003).
 CC -1- SIMILARITY: Belongs to the ATP-dependent AMP-binding enzyme
 CC family.
 CC EMBL; AB070951; BAB69332.1; --
 CC EMBL; AP005036; BAC71355.1; --
 CC HSP; O30409; IDNY.
 CC GO; GO:0003824; F:catalytic activity; IEA.
 CC GO; GO:0048037; F:cofactor binding; IEA.
 CC GO; GO:0008152; P:metabolism; IEA.
 CC InterPro; IPR010071; AA adenyl_dom.
 CC InterPro; IPR009081; ACP like.
 CC InterPro; IPR008073; AMP-bind.
 CC InterPro; IPR001242; Condensatn.
 CC InterPro; IPR000977; DNA_ligase.
 CC InterPro; IPR006163; Phosphateth_bind.
 CC InterPro; IPR006162; Phosphateth_bind.
 CC Pfam; PF00501; AMP-binding; 2.
 CC Pfam; PF00668; Condensation; 1.
 CC Pfam; PF00550; PP-binding; 2.
 CC PRINTS; PR00154; AMPBINDING.
 CC TIGRFBMS; TIGR01733; AA-adenyl-dom; 2.
 CC PROSITE; PS00075; ACP DOMAIN; 2.
 CC PROSITE; PS00697; DNA_LIGASE A1; UNKNOWN_1.
 CC PROSITE; PS00012; PHOSPHOPANTHETINE; UNKNOWN_1.
 CC Complete proteome.
 CC SEQUENCE 1709 AA; 191215 MW; 1E6A0D4D963A6783 CRC64;
 CC
 CC Query Match 3.6%; Score 9; DB 2; Length 1709;
 CC Best Local Similarity 100.0%; Pred. No. 18;
 CC Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC QY 229 PLGALELLS 237
 CC Db 1063 PLGALELLS 1071
 CC
 CC RESULT 3
 CC Q9YBB5 PRELIMINARY; PRT; 108 AA.
 CC ID Q9YBB5
 CC AC Q9YBB5
 CC DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 CC DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 CC DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 CC DE Hypothetical protein APE1682.
 CC GN OrderedLocusNames=APE1682;
 CC OS Aeropyrum pernix.
 CC OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcaceae;
 CC OC Desulfurococcaceae; Aeropyrum.
 CC OX NCBI_TaxID=56636;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN=K1;
 CC MEDLINE-9910339; PubMed=10382966;
 RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
 RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Kosugi H.,
 RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
 RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudo H. Y.,

RA Yamazaki J., Kushida N., Oguchi A., Aoki K.-I., Kubota K.,
 RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.,
 RT "Complete genome sequence of an aerobic hyper-thermophilic
 RT crenarchaeon, Aeropyrum pernix K1.",
 RL DNA Res. 6:83-101(1999)
 DR EMBL; AP000062; BAA80683.1; -.
 DR F1R; F72549; F72549.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 108 AA; 11640 MW; 073BBD9F5EE95106 CRC64;

Query Match 3.2%; Score 8; DB 2; Length 108;
 Best Local Similarity 100.0%; Pred. No. 16;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 217 TPRPLALV 224
 |||||
 DB 19 TPRPLALV 26

RESULT 4
 Q8R155 PRELIMINARY; PRT; 124 AA.
 AC Q8R155;
 DT 01-JUN-2002 (TRENBLrel. 21, Created)
 DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)
 DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
 DE CDNA sequence BC025446.
 GN Name=BC025446;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-FVB/N; TISSUE=Liver;
 RX MEDLINE=23388257; PubMed=1477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smalish D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RP SEQUENCE FROM N.A.
 RC STRAIN-FVB/N; TISSUE=Liver;
 RA Strausberg R.;
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC025446; AAB25446.1; -.
 DR MGI; MGI:2385015; BC025446.
 DR InterPro; IPR003632; LY-6 CD59.
 DR InterPro; IPR001526; LY6 UPAR.
 DR Pfam; PF00021; UPAR_LY6_1.
 DR ProDom; PD003128; LY-6 CD59; 1.
 DR SMART; SM00134; LU; 1.
 DR PROSITE; PS00983; LY6 UPAR; 1.
 SQ SEQUENCE 124 AA; 13401 MW; C53338BD584177B0 CRC64;

Query Match 3.2%; Score 8; DB 2; Length 124;
 Best Local Similarity 100.0%; Pred. No. 18;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 156 VFILLVAL 163
 |||||
 DB 11 VFILLVAL 18

RESULT 5
 Q6Y402 PRELIMINARY; PRT; 159 AA.
 AC Q6Y402;
 DT 05-JUL-2004 (TRENBLrel. 27, Created)
 DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)
 DE NADH dehydrogenase subunit 6.
 GN Name=nad6;
 OS Ophiura lutea.
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Asterozoa;
 OC Ophiuroidea; Ophiurida; Ophiurida; Ophiurina; Chilophiurina;
 OC Ophiuridae; Ophiura.
 OX NCBI_TaxID=224378;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX PubMed=15019608; DOI=10.1016/j.ympev.2003.07.005;
 RA Scouras A., Beckenbach K., Arndt A., Smith M.J.;
 RT "Complete mitochondrial genome DNA sequence for two ophiuroids and a
 RT holothuroid: the utility of protein gene sequence and gene maps in the
 RT analyses of deep deuterostome phylogeny.";
 RL Mol. Phylogenet. Evol. 31:50-65(2004).
 DR EMBL; AY184223; AAO65637.1; -.
 DR GO; GO:0005739; C:mitochondrion; IEA.
 DR GO; GO:0008137; F:NADH dehydrogenase (ubiquinone) activity; IEA.
 DR GO; GO:0006120; P:mitochondrial electron transport, NADH to u...; IEA.
 DR InterPro; IPR001457; Oxidored_q3.
 DR Pfam; PF00499; Oxidored_q3; 1.
 KW Mitochondrion.
 SQ SEQUENCE 159 AA; 17502 MW; 2758C60783FF5E86 CRC64;

Query Match 3.2%; Score 8; DB 2; Length 159;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 157 FILLVALV 164
 |||||
 DB 135 FILLVALV 142

RESULT 6
 Q8X1D5 PRELIMINARY; PRT; 163 AA.
 AC Q8X1D5;
 DT 01-MAR-2002 (TRENBLrel. 20, Created)
 DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
 DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
 DE Mig2-4.
 GN Name=mig2-4;
 OS Ustilago maydis (Smut fungus).
 OC Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes;
 OC Ustilaginomycetidae; Ustilaginales; Ustilaginaceae; Ustilago.
 OX NCBI_TaxID=5270;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21839640; PubMed=11849538;
 RA Basse C.W., Kolb S., Kahmann R.;
 RT "A maize-specifically expressed gene cluster in Ustilago maydis.";
 RL Mol. Microbiol. 43:75-93(2002).
 DR EMBL; AF429982; AAL67330.1; -.
 SQ SEQUENCE 163 AA; 17723 MW; 97A653FFA3E1E16 CRC64;

Query Match 3.2%; Score 8; DB 2; Length 163;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 LFLAASLS 27
| | | | | | | |
Db 7 LFLAASLS 14

RESULT 7
Q6AYS0 PRELIMINARY; PRT; 201 AA.

AC Q6AYS0
DT 25-OCT-2004 (TEMBLrel. 28, Created)
DT 25-OCT-2004 (TEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TEMBLrel. 28, Last annotation update)
DE Hypothetical protein.
DE Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OC NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heien F.,
RA Datschenko L., Marusina K., Farmer A.F., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedini T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McWean P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grinwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Kryzyski M.I., Skalski J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Director WGC Project;
RA Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC078937; AAH78937.1; -
DR InterPro; IPR007110; IG-like.
KW Hypothetical protein.
SQ SEQUENCE 201 AA; 22762 MW; 790B959BA4839B98 CRC64;

Query Match 3.2%; Score 8; DB 2; Length 201;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 52 VMSNCISN 59
| | | | | | | |
Db 51 VMSNCISN 58

RESULT 8
Q7NUR2 PRELIMINARY; PRT; 221 AA.

AC Q7NUR2
DT 01-MAR-2004 (TEMBLrel. 26, Created)
DT 01-MAR-2004 (TEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TEMBLrel. 26, Last annotation update)
DE Probable two-component response regulator.
GN OrderedLocusNames=CV2635;
OC Chromobacterium violaceum.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;

2

OX NCBI_TaxID=12230;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NBXC2;
 RA Shi M., Zhou X.;
 RT "Genomic variation of turnip mosaic virus.";
 RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ831813; CAH42011.1; -;
 KW Coat protein; Polyprotein.
 FT NON_TER 1
 FT CHAIN <1> 288 coat protein.
 FT NON_TER 288
 SQ SEQUENCE 288 AA; 33104 MW; AA7364110079E314 CRC64;
 Query Match 3.2%; Score 8; DB 2; Length 288;
 Best Local Similarity 100.0%; Pred. No. 38;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 188 QMKVAALR 195
 DB 240 QMKVAALR 247
 RESULT 10
 Q64WT9 PRELIMINARY; PRT; 304 AA.
 ID Q64WT9;
 AC Q64WT9;
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Hypothetical protein.
 GN ORFNames=BF1287;
 OS Bacteroides fragilis.
 OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;
 OC Bacteroidaceae; Bacteroides.
 OX NCBI_TaxID=817;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=YCH46;
 RA Kuwahara T., Yamashita A., Hirakawa H., Nakayama H., Toh H., Okada N.,
 RA Kuwara S., Hattori M., Hayashi T., Ohnishi Y.;
 RT "Genomic analysis of Bacteroides fragilis reveals extensive DNA
 RT inversions regulating cell surface adaptation.";
 RL Proc. Natl. Acad. Sci. U.S.A. 0:0-0(2004).
 DR EMBL; AP006841; BAD48037.1; -;
 KW Hypothetical protein.
 SQ SEQUENCE 304 AA; 33071 MW; 575A6D8FC7BEDAFA CRC64;
 Query Match 3.2%; Score 8; DB 2; Length 304;
 Best Local Similarity 100.0%; Pred. No. 40;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 20 LFLAASLS 27
 DB 124 LFLAASLS 131
 RESULT 11
 Q6P2X0 PRELIMINARY; PRT; 316 AA.
 ID Q6P2X0;
 AC Q6P2X0;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Hypothetical protein MGC76280.
 GN Name=MGC76280;
 OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8364;
 RN [1]
 RP SEQUENCE FROM N.A.

RC TISSUE=Embryo;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Mullighy S.J.,
 RA Raba S.S., Lequellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huiyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettner M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RL Klein S., Gerhard D.S.;
 RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RL Klein S., Gerhard D.S.;
 RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC064266; AAH64266.1; -;
 DR EMBL; BC080950; AAH80950.1; -;
 DR GO; GO:0004519; F:endonuclease activity; IEA.
 DR GO; GO:0006281; P:DNA repair; IEA.
 DR InterPro; IPR000097; ApEndonclisel.
 DR InterPro; IPR004808; ExoIII_xth.
 DR InterPro; IPR005135; Exo endo phos.
 DR Pfam; PF03372; Exo endo phos; 1.
 DR TIGRFAMs; TIGR00195; exoDNase_III; 1.
 DR TIGRFAMs; TIGR00633; xth; 1.
 DR PROSITE; PS00728; AP_NUCLEASE_F1_3; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 316 AA; 35955 MW; AB4E25B5215497B1 CRC64;
 Query Match 3.2%; Score 8; DB 2; Length 316;
 Best Local Similarity 100.0%; Pred. No. 42;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 132 EVSGAEPQ 139
 DB 9 EVSGAEPQ 16
 RESULT 12
 Q7Y5I6 PRELIMINARY; PRT; 332 AA.
 ID Q7Y5I6;
 AC Q7Y5I6;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE 31L.
 OS Xanthomonas oryzae bacteriophage Xp10.
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales.
 OX NCBI_TaxID=232237;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22734642; PubMed=12850143; DOI=10.1016/S0022-2836(03)00634-X;
 RA Yuzenkova J., Nechaev S., Berlin J., Rogulja D., Kuznedelov K.,
 RA Innan R., Mushagian A., Severinov K.;
 RT "Genome of Xanthomonas oryzae bacteriophage Xp10: an odd T-odd
 RT phase.";

```

RL J. Mol. Biol. 330:735-748(2003).
RN [2]
RP SEQUENCE FROM N.A.
RA Nechaev S.Y., Yuzenkova Y.V., Berlin J., Regulja D.; Kurnedelov K.,
RA Inman R., Mushhegian A., Severinov K.V.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY298121; AAP58698.1; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003910; F:DNA ligase (ATP) activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0006310; P:DNA recombination; IEA.
DR GO; GO:0006281; P:DNA repair; IEA.
DR GO; GO:0006260; P:DNA replication; IEA.
DR InterPro; IPR000977; DNA ligase.
DR InterPro; IPR008994; Nucleic acid OB.
DR Pfam; PF01068; DNA_ligase_A; 1.
SQ SEQUENCE 332 AA; 36551 MW; 6CB6C37F945D6CBF CRC64;

Query Match 3.2%; Score 8; DB 2; Length 332;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 LAASLSAQ 29
DB 193 LAASLSAQ 200
|||||
|||||

RESULT 13
OB3402 PRELIMINARY; PRT; 384 AA.
AC OB3402;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Cell division protein (FtsW).
GN OrderedLocNames=FP0387;
OS Treponema pallidum.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=160;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=Nichols;
RX MEDLINE=98332770; PubMed=9665876; DOI=10.1126/science.281.5375.375;
RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
RA Dodson R.J., Gwinn M.L., Hickey E.K., Clayton R.A., Ketchum K.A.,
RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S.L., Peterson J.D.,
RA Khalak H.G., Richardson D.L., Howell J.K., Chidambaram M.,
RA Utterback T.R., McDonald L.A., Artach P., Bowman C., Cotton M.D.,
RA Fujii C., Garland S.A., Hatch B., Horst K., Roberts K.M., Sandusky M.,
RA Weidman J.F., Smith H.O., Venter J.C.;
RA "Complete genome sequence of Treponema pallidum, the syphilis
RT spirochete."
RL Science 281:375-388(1998).
DR EMBL; AE001217; AAC65371.1; -.
DR PIR; E71331; E71331..
DR TIGR; TP0387; -.
DR GO; GO:0006021; C:integral to membrane; IEA.
DR GO; GO:0007049; F:cell cycle; IEA.
DR InterPro; IPR001182; Cell_Cycle.
DR InterPro; IPR006162; Ppantne S.
DR Pfam; PF01098; FTSW RODA SPOVE; 1.
DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 384 AA; 42714 MW; 01DF27E5E30B4F2E CRC64;

Query Match 3.2%; Score 8; DB 2; Length 384;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 156 VFILLVAL 163
DB 19 VFILLVAL 26
|||||
|||||

us-10-080-522-1.oli.rup

RESULT 14
OB97Y6 PRELIMINARY; PRT; 521 AA.
AC OB97Y6;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Putative efflux pump component mtrP.
GN Names=mtrP; OrderedLocNames=CTC00584;
OS Clostridium tetani.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1513;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=Massachusetts / E88;
RX MEDLINE=22457253; PubMed=12552129; DOI=10.1073/pnas.0335853100;
RA Brueggemann H., Baeumer S., Fricke W.F., Wierzer A., Liesegang H.,
RA Decker I., Herzberg C., Martinez-Arias R., Merkl R., Henne A.,
RA Gottschalk G.;
RA "The genome sequence of Clostridium tetani, the causative agent of
RT tetanus disease."
RL Proc. Natl. Acad. Sci. U.S.A. 100:1316-1321(2003).
DR EMBL; AE015938; AAC035200.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR004697; Put transp AbgT.
DR Pfam; PF03806; ABG_transport; 1.
KW Complete proteome.
SQ SEQUENCE 521 AA; 55829 MW; D1A9096671E58C8C CRC64;

Query Match 3.2%; Score 8; DB 2; Length 521;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 157 FILLVALV 164
DB 400 FILLVALV 407
|||||
|||||

RESULT 15
OB2FX6 PRELIMINARY; PRT; 545 AA.
AC OB2FX6;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Putative membrane protein.
GN OrderedLocNames=SAV4126;
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=NA-4680;
RX MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.211433198;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Oonoe T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RA "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites."
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN=NA-4680;
RX MEDLINE=22608306; PubMed=12692562;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RA "Complete genome sequence and comparative analysis of the industrial
RT
```

RT microorganism Streptomyces avermitilis."
RL Nat. Biotechnol. 21:526-531(2003).
DR EMBL; AP005037; BAC71838.1; -
DR GO: GO:0016787; P:hydrolase activity; IEA.
DR InterPro; IPR004843; M-pesterase.
DR Pfam; PF00149; Metallophos; 1.
KW Complete proteome.
SQ SEQUENCE 545 AA; 57993 MW; BDCB0CA12396479 CRC64;

Query Match 3.2%; Score 8; DB 2; Length 545;
Best Local Similarity 100.0%; Pred.No. 67;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 SPLGALEL 235
|||
Db 120 SPLGALEL 127

Search completed: March 20, 2005, 10:18:21
Job time : 96 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 20, 2005, 06:49:53 ; Search time 85 Seconds
(without alignments)
1128.430 Million cell updates/sec

Title: US-10-080-522-1

Perfect score: 1308

Sequence: 1 MQTCPLAPGHVSQALGTL.....PLGALELISQPLFPFYAADP 248

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

- 1: geneseqp1980s:*
- 2: geneseqp1990s:*
- 3: geneseqp2000s:*
- 4: geneseqp2001s:*
- 5: geneseqp2002s:*
- 6: geneseqp2003as:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1308	100.0	248	2	Aaw48811 K12 prote
2	1308	100.0	248	2	Aay21846 Human sig
3	1308	100.0	248	4	Aab36658 Human K12
4	1308	100.0	248	8	Adj75366 Marker ge
5	1308	100.0	248	8	Adp24664 PRO polyp
6	916.5	70.1	183	6	Aao29897 Human org
7	720.5	55.1	149	7	Adb36336 Human imm
8	719.5	55.0	162	6	Aao29896 Human org
9	492	37.6	101	6	Aao29895 Human org
10	326	24.9	212	4	Aab36660 Mouse K12
11	326	24.9	212	8	Adj76167 Marker ge
12	325	24.8	212	8	Adj76168 Marker ge
13	294	22.5	107	6	Aao29898 Human org
14	286	21.9	55	3	Aab34671 Gene 35 h
15	279	21.3	55	3	Aab34672 Human sec
16	163	12.5	30	3	Aab34673 Gene 35 h
17	156	11.9	43	7	Adb36337 Human imm
18	100	7.6	874	4	ABBS9013 Drosophil
19	96	7.3	435	6	ABR41273 Human DIT
20	96	7.3	442	4	Aab95569 Human pro
21	95.5	7.3	506	5	Adi16910 Cow NOVX
22	91	7.0	6620	7	Adj70485 Human hea
23	91	7.0	7968	6	ABG76187 Human ser
24	90.5	6.9	1171	7	ABO63781 Klebsiell
25	89	6.8	435	4	AAB61190 Mature hu

ALIGNMENTS

RESULT 1

AAW48811
ID AAW48811 standard; protein; 248 AA.

XX AC AAW48811;

DT 26-OCT-1998 (first entry)

XX DE K12 protein.

XX KW DNA probe; CD7 Hs1 DNase hypersensitive site; mRNA northern blot;

XX KW human erythroleukemic; HEL; K562 cell line; clone; breast cancer;

XX KW ovarian cancer; malignant; tumour.

OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Misc-difference 239

FT FT /note= "encoded by CCAA"

XX PN WO9822502-A1.

XX PD 28-MAY-1998.

XX PF 24-NOV-1997; 97WO-US021517.

XX PR 22-NOV-1996; 96US-00755559.

XX FA (UYDU-) UNIV DUKE.

XX PI Kaufman RE, Slentz-Kesler KA;

XX DR WPI; 1998-312415/27.

XX DR N-PSDB; AAV32446.

XX PT New isolated K12 protein gene - which is over expressed in certain

XX PT neoplastic cells, used to develop products for tumour detection and

XX PS Claim 3; Fig 1; 44pp; English.

XX CC This present sequence represents the K12 protein, the gene for which has

XX CC been located on chromosome 17q25. To obtain this genes cDNA sequence a

XX CC 500 bp DNA probe, which can be located just upstream of the CD7 Hs1 DNase

XX CC hypersensitive site, was used against a mRNA northern blot. From this a

XX CC 1.8kb transcript was detected in the human erythroleukemic cell line HEL.

XX CC The probe was then used to screen a human erythroleukemic cell line K562

XX CC cDNA library, from which several clones were identified and isolated that

CC constituted a 1.8kb cDNA. This cDNA was designated K12 and was found to
 CC have a single open reading frame as well as being in the same orientation
 CC as CD7. The K12 gene was found to be expressed in both breast and ovarian
 CC cancer cells at a much higher level than any other malignant or normal
 CC tissue that was examined, thus enabling the K12 to be a useful protein in
 CC tumour detection and treatment

XX
 SQ Sequence 248 AA;
 Query Match 100.0%; Score 1308; DB 2; Length 248;
 Best Local Similarity 100.0%; Pred. No. 1.8e-133;
 Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MQTCLAPFGHVSQALGTLFLAASLSAQNEGWDSPTCTEGVSVSGENTVMSCNISNA 60
 DB 1 MQTCLAPFGHVSQALGTLFLAASLSAQNEGWDSPTCTEGVSVSGENTVMSCNISNA 60
 QY 61 FSHVNIKLRAHQESAI FNEVAPGYFSRDGWLQVGGVAQLVKGARDSHAGLYMHLV 120
 DB 61 FSHVNIKLRAHQESAI FNEVAPGYFSRDGWLQVGGVAQLVKGARDSHAGLYMHLV 120
 QY 121 GHQRNRQVTLVSGAEPOSADPTGFWPVAVTAVFILLVALVNFAYRCRCQORREK 180
 DB 121 GHQRNRQVTLVSGAEPOSADPTGFWPVAVTAVFILLVALVNFAYRCRCQORREK 180
 QY 181 KFFLEPQMKVAALRAGAQQGLSRASAEIWTDPSEPTPRPLALVFKPSPLGALLELLSPQ 240
 DB 181 KFFLEPQMKVAALRAGAQQGLSRASAEIWTDPSEPTPRPLALVFKPSPLGALLELLSPQ 240
 QY 241 LFPYAADP 248
 DB 241 LFPYAADP 248

RESULT 2
 AAY21846
 ID AAY21846 standard; protein; 248 AA.
 XX AAY21846;
 XX 20-SEP-1999 (first entry)
 XX Human signal peptide-containing protein (SIGP) (clone ID 1747327).
 XX Signal-peptide containing protein; SIGP; human; cancer; immune response;
 KW adenocarcinoma; leukemia; lymphoma; melanoma; myeloma; sarcoma; AIDS;
 KW Addison's disease; adult respiratory distress syndrome; allergy; anemia;
 KW asthma; atherosclerosis; bronchitis; cholelactitis; Crohn's disease;
 KW ulcerative colitis; atopic dermatitis; dermatomyositis; emphysema;
 KW diabetes mellitus; atrophic gastritis; glomerulonephritis; gout; trauma;
 KW Grave's disease; hyperesoinophilia; irritable bowel syndrome; infection;
 KW lupus erythematosus; multiple sclerosis; myasthenia gravis; inflammation;
 KW osteoarthritis; osteoporosis; pancreatitis; polymyositis; scleroderma;
 KW rheumatoid arthritis; Sjogren's syndrome; autoimmune thyroiditis.
 XX Homo sapiens.
 OS
 XX WO9933981-A2.
 XX 08-JUL-1999.
 XX 22-DEC-1998; 98WO-US027598.
 XX 31-DEC-1997; 97US-00002485.
 XX (INCY-) INCYTE PHARM INC.
 XX Lal P, Hillman JL, Corley NC, Guegler KJ, Baughn MR, Sather SK;
 PI Shah P;
 XX WPI, 1999-430242/36.
 DR N-PSDB; AAX82081.
 XX

PT Human signal-peptide containing protein coding sequences used to treat
 PT cancer and immune responses.

XX Claim 1; Page 79-80; 99pp; English.

XX The invention provides human signal-peptide containing proteins (SIGP)
 CC (AAY21841-855) and polynucleotides (AAX82076-90) encoding the proteins. A
 CC host cell containing a vector comprising SIGP DNA can be used to produce
 CC the SIGP protein. The SIGP protein can be used, in conjunction with a
 CC pharmaceutical carrier to treat or prevent a cancer. An antagonist of the
 CC SIGP protein can be used to treat or prevent a cancer or an immune
 CC response. The cancers that can be treated or prevented include sarcomas,
 CC adenocarcinomas, leukemia's, lymphomas, melanomas, teratocarcinomas,
 CC myelomas and cancers of the adrenal gland, bladder, bone, bone marrow,
 CC brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract,
 CC heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis,
 CC prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and
 CC uterus. The immune responses that can be treated or prevented include,
 CC AIDS, Addison's disease, adult respiratory distress syndrome, allergies,
 CC anemia, asthma, atherosclerosis, bronchitis, cholelactitis, Crohn's
 CC disease, ulcerative colitis, atopic dermatitis, dermatomyositis, diabetes
 CC mellitus, emphysema, atrophic gastritis, glomerulonephritis, Grave's
 CC disease, gout, hyperesoinophilia, irritable bowel syndrome, lupus
 CC erythematosus, multiple sclerosis, myasthenia gravis, myocardial or
 CC pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis,
 CC polymyositis, rheumatoid arthritis, scleroderma, Sjogren's syndrome, and
 CC autoimmune thyroiditis, complications of cancer, infections, and trauma

XX Sequence 248 AA;

Query Match 100.0%; Score 1308; DB 2; Length 248;
 Best Local Similarity 100.0%; Pred. No. 1.8e-133;
 Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MQTCLAPFGHVSQALGTLFLAASLSAQNEGWDSPTCTEGVSVSGENTVMSCNISNA 60
 DB 1 MQTCLAPFGHVSQALGTLFLAASLSAQNEGWDSPTCTEGVSVSGENTVMSCNISNA 60
 QY 61 FSHVNIKLRAHQESAI FNEVAPGYFSRDGWLQVGGVAQLVKGARDSHAGLYMHLV 120
 DB 61 FSHVNIKLRAHQESAI FNEVAPGYFSRDGWLQVGGVAQLVKGARDSHAGLYMHLV 120
 QY 121 GHQRNRQVTLVSGAEPOSADPTGFWPVAVTAVFILLVALVNFAYRCRCQORREK 180
 DB 121 GHQRNRQVTLVSGAEPOSADPTGFWPVAVTAVFILLVALVNFAYRCRCQORREK 180
 QY 181 KFFLEPQMKVAALRAGAQQGLSRASAEIWTDPSEPTPRPLALVFKPSPLGALLELLSPQ 240
 DB 181 KFFLEPQMKVAALRAGAQQGLSRASAEIWTDPSEPTPRPLALVFKPSPLGALLELLSPQ 240
 QY 241 LFPYAADP 248
 DB 241 LFPYAADP 248

RESULT 3
 AAB36658
 ID AAB36658 standard; protein; 248 AA.
 XX AAB36658;
 XX 13-MAR-2001 (first entry)
 XX Human K12 protein sequence SEQ ID NO:4.

XX Human; CD7; K12; cognate ligand; cluster of differentiation; cancer;
 KW identification; inhibiting T cell proliferation; HIV; infection;
 KW activating natural killer cell proliferation; leukaemia; lymphoma;
 KW sepsis; graft versus host disease; autoimmune disease; arthritis;
 KW multiple sclerosis; rheumatoid arthritis; psoriatic arthritis; lupus;
 KW scleroderma; psoriasis; atopic dermatitis; type 1 diabetes mellitus;
 KW Hashimoto's thyroiditis; pernicious anaemia; Addison's disease; uveitis;
 KW myasthenia gravis; psoriasis; Guillain-Barre Syndrome; Grave's disease;

KW systemic lupus erythematosus; dermatomyositis; asthma; eczema;
KW atopic dermatitis; contact dermatitis; eczematous dermatitis;
KW seborrheic dermatitis; rhinitis.

XX Homo sapiens.

XX WO200073333-A2.

XX 07-DEC-2000.

XX 26-MAY-2000; 2000WO-US014612.

XX 28-MAY-1999; 99US-0136450P.

XX (IMMV) IMMUNEX CORP.

XX Lyman SD, Fanslow WC;

XX WPI; 2001-061511/07.

XX N-PSDB; AAC88152.

XX Stimulating intracellular signaling of CD7 comprises contacting a cell
PT expressing CD7 with recombinant K12 protein, the cognate ligand of CD7,
PT to inhibit T cell proliferation and/or activate natural killer cell
PT proliferation.

XX Claim 2; Page 38-39; 42pp; English.

XX The present invention describes a method for stimulating (S) the
CC intracellular signalling of CD (Cluster of differentiation) 7 comprising
CC contacting a cell that expresses CD7 with a recombinant K12 protein (I),
CC the cognate ligand of CD7. (S) is useful for inhibiting T cell
CC proliferation and/or activating NK (natural killer) cell proliferation
CC and/or inducing NK toxicity in a mammal which involves administration of
CC K12 protein. It is also used for treating HIV-1 infection, cancer (T cell
CC leukaemia, acute lymphocytic leukaemia, cutaneous T cell lymphoma),
CC bacterial and viral infections, mediated by CD7. In the case of treating
CC T cell leukaemia the soluble K12 protein is covalently attached to a
CC toxin. A disease mediated by CD7 such as sepsis, graft versus host
CC disease due to transplantation, autoimmune diseases, multiple sclerosis,
CC arthritis, rheumatoid arthritis, psoriatic arthritis, scleroderma, lupus,
CC psoriasis, atopic dermatitis, type 1 diabetes mellitus, Hashimoto's
CC thyroiditis, pernicious anaemia, Addison's disease, myasthenia gravis,
CC uveitis, psoriasis, Guillain-Barre Syndrome, Grave's disease, systemic
CC lupus erythematosus and dermatomyositis, asthma, eczema, atopic
CC dermatitis, contact dermatitis, other eczematous dermatides, seborrheic
CC dermatitis, and rhinitis is also treated by administering a K12
CC antagonist (neutralising antibody). The present sequence represents the
CC human K12 protein, which is given in the exemplification of the present
XX invention

XX Sequence 248 AA;

Query Match 100.0%; Score 1308; DB 4; Length 248;
Best Local Similarity 100.0%; Pred. No. 1.8e-133;
Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MOTCLAPGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSNCISNA 60
DB 1 MOTCLAPGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSNCISNA 60
QY 61 FSHVNIKLRAHQESAIENEVAPGYFSRDGWLQVGGVAQLVIKGDARDSHAGLYMHLV 120
DB 61 FSHVNIKLRAHQESAIENEVAPGYFSRDGWLQVGGVAQLVIKGDARDSHAGLYMHLV 120
QY 121 GQRRNRQVTLVSGAEPQAPDPTGFWPVAVTAVFILLVALNMFAYWRCRCQORREK 180
DB 121 GQRRNRQVTLVSGAEPQAPDPTGFWPVAVTAVFILLVALNMFAYWRCRCQORREK 180
QY 181 KFFLEPQMKVAALRAGAQOGLSRASAEIWTDPDSEPTPRPLALVFKPSPLGALLSPQ 240
DB 181 KFFLEPQMKVAALRAGAQOGLSRASAEIWTDPDSEPTPRPLALVFKPSPLGALLSPQ 240

QY 241 LFPYAADP 248
DB 241 LFPYAADP 248

RESULT 4
ADJ75366

XX ADJ75366 standard; protein; 248 AA.

XX AC ADJ75366;

XX 20-MAY-2004 (first entry)

XX Marker gene related amino acid sequence SEQ ID NO:618.

XX bronchial asthma; chronic obstructive pulmonary disease;
KW respiratory epithelial cell; interleukin-13; respiratory; antiasthmatic;
KW gene therapy; marker.

XX Homo sapiens.

XX EPI394274-A2.

XX 03-MAR-2004.

XX 04-AUG-2003; 2003EP-00254857.

XX 06-AUG-2002; 2002JP-00229312.

XX 20-MAR-2003; 2003JP-00077212.

XX (GENO-) GENOX RES INC.

XX Ohtani N, Sugita Y, Yamaya M, Kubo H, Nagai H, Izuhara K;

XX WPI; 2004-193155/19.

XX Testing for bronchial asthma or chronic obstructive pulmonary disease by
PT comparing the expression level of a marker gene in a biological sample
PT from a subject with the expression level of the gene in a sample from a
PT healthy subject.

XX Example 11; SEQ ID NO 618; 241pp; English.

XX The present invention describes a method of testing for bronchial asthma
or chronic obstructive pulmonary disease. The method comprises
CC determining the expression level of a marker gene in a biological sample
CC from a subject, comparing the expression level determined with the
CC expression level of the marker gene in a biological sample from a healthy
CC subject, and judging whether the subject has bronchial asthma or chronic
CC obstructive pulmonary disease. The marker gene comprises: (a) a group of
CC genes (S1) whose expression levels increase when respiratory epithelial
CC cells are stimulated with interleukin-13; or (b) a group of genes (S2)
CC whose expression levels decrease when respiratory epithelial cells are
CC stimulated with interleukin-13. Also described: (1) a reagent (I) for
CC testing for bronchial asthma or chronic obstructive pulmonary disease;
CC (2) a kit for screening for a candidate compound for a therapeutic agent
CC to treat bronchial asthma or chronic obstructive pulmonary disease; (3)
CC an animal model for bronchial asthma or chronic obstructive pulmonary
CC disease; (4) an inducer that induces bronchial asthma in a mouse; (5) a
CC method for producing an animal model for bronchial asthma or chronic
CC obstructive pulmonary disease; (6) a therapeutic agent for bronchial
CC asthma or chronic obstructive pulmonary disease, comprising the compound,
CC the marker gene or an antisense nucleic acid corresponding to a portion of
CC the marker gene, a ribozyme, a polynucleotide that suppresses the
CC expression of the gene through an RNAi effect or an antibody recognising
CC a protein encoded by a marker gene; and (7) a DNA chip for testing for
CC bronchial asthma or a chronic obstructive pulmonary disease, on which a
CC probe has been immobilised to assay a marker gene. (I) has respiratory
CC and antiasthmatic activities, and can be used in gene therapy. The method
CC is useful for testing for or screening for a therapeutic agent for
CC bronchial asthma or chronic obstructive pulmonary disease. The present
XX sequence is used in the exemplification of the present invention.

SQ Sequence 248 AA;
 Query Match 100.0%; Score 1308; DB 8; Length 248;
 Best Local Similarity 100.0%; Pred. No. 1.8e-133;
 Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MOTCTPLAFPGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSCNISNA 60
 DB 1 MOTCTPLAFPGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSCNISNA 60

QY 61 FSHVNIKLRAHQESAI FNEVAPGYFSRDGWLQVGGVAQLVKGARDSHAGLYMHLV 120
 DB 61 FSHVNIKLRAHQESAI FNEVAPGYFSRDGWLQVGGVAQLVKGARDSHAGLYMHLV 120

QY 121 GHQRNNQVTLVSGAEPQSDPTGFWPVPVAVTAVFILLVALVMFAWYRCRCQORREK 180
 DB 121 GHQRNNQVTLVSGAEPQSDPTGFWPVPVAVTAVFILLVALVMFAWYRCRCQORREK 180

QY 181 KFFLLEPQMKVAALRAGAQQGLSRASAEI WTPDSEPTPRPLALVFKPSPLGALLELLSPQ 240
 DB 181 KFFLLEPQMKVAALRAGAQQGLSRASAEI WTPDSEPTPRPLALVFKPSPLGALLELLSPQ 240

QY 241 LFPYAADP 248
 DB 241 LFPYAADP 248

RESULT 5
 ADP24664
 ID ADP24664 standard; protein; 248 AA.
 XX
 AC ADP24664;
 DT 18-NOV-2004 (first entry)
 DE PRO polypeptide SEQ ID NO:1842.
 KW PRO; antiinflammatory; antiarthritic; antirheumatic; immunosuppressive;
 KW osteopathic; antidiabetic; dermatological; antipsoriatic; antiallergic;
 KW antiasthmatic; hepatotropic; respiratory; gene therapy; immune system.
 XX Unidentified.
 XX WO2004041170-A2.
 XX 21-MAY-2004.
 XX 30-OCT-2003; 2003WO-US034312.
 XX 01-NOV-2002; 2002US-0423394P.
 XX (GETH) GENENTECH INC.
 XX Clark H, Schoenfeld J, Van Lookeren M, Williams PM, Wood WI;
 XX Wu TD;
 XX WPI; 2004-419628/39.
 XX N-PSDB; ADP24663.
 XX New PRO polypeptides and polynucleotides, useful for treating e.g.
 PT erythematous, rheumatoid arthritis, diabetes mellitus, immune-mediated
 PT renal disease, or demyelinating diseases of the central or peripheral
 PT nervous system.
 XX Claim 7; SEQ ID NO 1842; 2940pp; English.
 PS The invention relates to a novel isolated nucleic acid and the PRO
 CC polypeptide encoded by it. A protein of the invention has
 CC antiinflammatory, antiarthritic, antirheumatic, immunosuppressive,
 CC osteopathic, antidiabetic, dermatological, antipsoriatic, antiallergic,
 CC antiasthmatic, hepatotropic, and respiratory activity. A polynucleotide
 CC of the invention may have a use in gene therapy. The PRO polypeptide, its
 CC agonist, antagonist, or antibody that specifically binds to the

CC polypeptide is useful for treating an immune related disorder such as
 CC systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis,
 CC juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis, an
 CC idiopathic inflammatory myopathy, Sjogren's syndrome, autoimmune
 CC vasculitis, sarcoidosis, autoimmune haemolytic anaemia, systemic
 CC thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal
 CC disease, a demyelinating disease of the central or peripheral nervous
 CC system, idiopathic demyelinating polyneuropathy, Guillain-Barre syndrome,
 CC a chronic inflammatory demyelinating polynuropathy, a hepatobiliary
 CC disease, infectious or autoimmune chronic active hepatitis, primary
 CC biliary cirrhosis, granulomatous hepatitis, sclerosing cholangitis,
 CC inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's
 CC disease, an autoimmune or immune-mediated skin disease, a bullous skin
 CC disease, erythema multiforme, contact dermatitis, psoriasis, an allergic
 CC disease, asthma, allergic rhinitis, atopic dermatitis, food
 CC hypersensitivity, urticaria, an immunologic disease of the lung,
 CC eosinophilic pneumonia, idiopathic pulmonary fibrosis, hypersensitivity
 CC pneumonitis, a transplantation associated disease, graft rejection or
 CC graft-versus-host disease. The present sequence represents a PRO protein
 CC of the invention.
 XX
 XX SQ Sequence 248 AA;

Query Match 100.0%; Score 1308; DB 8; Length 248;
 Best Local Similarity 100.0%; Pred. No. 1.8e-133;
 Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MOTCTPLAFPGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSCNISNA 60
 DB 1 MOTCTPLAFPGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSCNISNA 60

QY 61 FSHVNIKLRAHQESAI FNEVAPGYFSRDGWLQVGGVAQLVKGARDSHAGLYMHLV 120
 DB 61 FSHVNIKLRAHQESAI FNEVAPGYFSRDGWLQVGGVAQLVKGARDSHAGLYMHLV 120

QY 121 GHQRNNQVTLVSGAEPQSDPTGFWPVPVAVTAVFILLVALVMFAWYRCRCQORREK 180
 DB 121 GHQRNNQVTLVSGAEPQSDPTGFWPVPVAVTAVFILLVALVMFAWYRCRCQORREK 180

QY 181 KFFLLEPQMKVAALRAGAQQGLSRASAEI WTPDSEPTPRPLALVFKPSPLGALLELLSPQ 240
 DB 181 KFFLLEPQMKVAALRAGAQQGLSRASAEI WTPDSEPTPRPLALVFKPSPLGALLELLSPQ 240

QY 241 LFPYAADP 248
 DB 241 LFPYAADP 248

RESULT 6
 AAO29897
 ID AAO29897 standard; protein; 183 AA.
 XX
 AC AAO29897;
 XX
 DT 03-SEP-2003 (first entry)
 DE Human organelle-associated protein (ORGA)-10.
 XX
 KW Human; organelle-associated protein; ORGA; cell proliferative disorder;
 KW actinic keratosis; arteriosclerosis; reproductive disorder; infertility;
 KW endometriosis; gastrointestinal disorder; ectopic pregnancy; hepatitis;
 KW gastroenteritis; neurological disorder; Alzheimer's disease; cirrhosis;
 KW glomerulonephritis; renal amyloidosis; renal failure; Addison's disease;
 KW renal disorder; ovulatory defect; teratogenesis; Pick's disease; cancer;
 KW Huntington's disease; urologic disorder; cystitis; dementia; dysphagia;
 KW indigestion; gastritis; hypertension; gene therapy; prostatic; anorexia;
 KW infection; urethritis; prostatitis.
 XX Homo sapiens.
 OS
 XX WO2003044171-A2.
 XX
 XX 30-MAY-2003.
 PD

```

XX PF 15-NOV-2002; 2002WO-US036807.
XX PR 16-NOV-2001; 2001US-0332384P.
XX PR 13-DEC-2001; 2001US-0341187P.
XX PR 23-JAN-2002; 2002US-0351151P.
XX PR 27-FEB-2002; 2002US-0360269P.
XX PR 05-APR-2002; 2002US-0370637P.
XX PR 14-JUN-2002; 2002US-0388946P.
XX PA (INCY-) INCYTE GENOMICS INC.
XX PI Ramkumar J, Burford N, Xu Y, Chawla NK, Kable AE, Becha SD;
PI Richardson TW, Gorvad AB, Sprague WM, Elliott VS, Khare R;
PI Hawkins PR, Jin P, Lee SY, Marquis JP, Swarnakar A, Wilson AD;
PI Hafalia AJA, Bulloch S;
XX WPI: 2003-457603/43.
XX DR N-PSDB; AAL60550.
XX PT New organelle-associated proteins and polynucleotides, useful for
XX PT diagnosing, treating and/or preventing cell proliferative, reproductive,
XX PT gastrointestinal, neurological, urologic, and renal disorders.
XX PS Claim 1; Page 166-167; 194pp; English.
XX CC The invention relates to human organelle-associated proteins (ORGA) and
XX CC nucleic acid molecules encoding such proteins. ORGA sequences are useful
XX CC for diagnosing, treating and/or preventing cell proliferative disorders
XX CC (e.g. actinic keratosis, arteriosclerosis, bursitis, cirrhosis, cancers
XX CC or hepatitis), reproductive disorders (e.g. infertility, endometriosis,
XX CC ovulatory defects, ectopic pregnancy or teratogenesis), gastrointestinal
XX CC disorders (e.g. dysphagia, peptide oesophagitis, gastritis, indigestion,
XX CC anorexia or gastroenteritis), neurological disorders (e.g. Alzheimer's
XX CC disease, Pick's disease, Huntington's disease or dementia), urologic
XX CC disorders (e.g. infection, urethritis, cystitis, prostatitis or upper
XX CC tract infections) and renal disorders (e.g. glomerulonephritis, renal
XX CC amyloidosis, renal failure, Addison's disease or hypertension). The
XX CC invention is also useful in gene therapy. The present sequence is human
XX CC ORGA protein
XX CC Sequence 183 AA;
XX SQ
Query Match 70.1%; Score 916.5; DB 6; Length 183;
Best Local Similarity 73.4%; Pred. No. 4.9e-91;
Matches 182; Conservative 0; Mismatches 1; Indels 65; Gaps 1;
QY 1 MOTCLAPFCHVSQALGTLFLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSCNINA 60
Db 1 MOTCLAPFCHVSQALGTLFLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSCNINA 60
QY 61 FSHVNIKLRAHQESAIFNEVAPGVFSDGWLQVGGVAQLVKGARDSHAGLYMHLV 120
Db 61 FSHVNIKLRAHQESAIFNEVAPGVFSDGWLQVGGVAQLVKGARDSHAGLYMHLV 120
QY 121 GHQRNRQVTLVSGAEPQSDPTDGFVPVAVTAVFLLVALVNFVAFWVRCRCQQORREK 180
Db 121 GH----- 122
QY 181 KFFLEPQKVAALRAGAQOGLSRASAEIWTDPDSFTPRPLALVPKPSPLGALELLSQP 240
Db 123 -----QTKVAALRAGAQOGLSRASAEIWTDPDSFTPRPLALVPKPSPLGALELLSQP 175
QY 241 LFPYAADP 248
Db 176 LFPYAADP 183

```

RESULT 7
ADB36336

ID ADB36336 standard; protein; 149 AA.

XX AC ADB36336;

```

XX DT 04-DEC-2003 (first entry)
XX XX Human immune response associated protein IRAP-7 SEQ ID NO:7.
XX DE human; immune response associated protein; IRAP; anti-HIV; anti-allergic;
XX KW antianemic; antiasthmatic; antiarteriosclerotic; dermatological;
XX KW antiinflammatory; antidiabetic; nephrotropic; antithyroid; thyromimetic;
XX KW immunosuppressive; antirheumatic; antiarthritic; osteopathic;
XX KW antibacterial; virucide; antiparasitic; protozoacide; fungicide;
XX KW cerebroprotective; neuroprotective; nootropic; antiparkinsonian;
XX KW antipsoriatic; cytostatic; cardiant; gene therapy;
XX KW immune system disorder; neurological disorder; developmental disorder;
XX KW muscle disorder; cell proliferative disorder; AIDS; allergy; anemia;
XX KW asthma; atherosclerosis; contact dermatitis; Crohn's disease; diabetes;
XX KW glomerulonephritis; Grave's disease; Hashimoto's thyroiditis;
XX KW multiple sclerosis; rheumatoid arthritis; osteoporosis;
XX KW systemic lupus erythematosus; infection; stroke; Alzheimer's disease;
XX KW Parkinson's disease; psoriasis; cancer; cardiomyopathy.
XX OS Homo sapiens.
XX XX WO2003074726-A2.
XX PN 12-SEP-2003.
XX PD 28-FEB-2003; 2003WO-US006307.
XX PF 01-MAR-2002; 2002US-0361088P.
XX PR 27-MAR-2002; 2002US-0368494P.
XX PR 10-MAY-2002; 2002US-0379876P.
XX PR 28-JUN-2002; 2002US-0392641P.
XX PA (INCY-) INCYTE GENOMICS INC.
XX PI Chawla NK, Becha SD, Jin P, Tran UK, Kable AE, Lee S;
XX PI Marquis JP, Emerling BM, Ison CH, Ramkumar J, Chang H;
XX WPI: 2003-722079/68.
XX DR N-PSDB; ADB36352.
XX XX New human immune response associated proteins and polynucleotides, useful
XX PT for diagnosing, preventing or treating diseases or conditions associated
XX PT with aberrant enzyme expression, e.g. cancer, AIDS, atherosclerosis or
XX PT stroke.
XX PS Claim 1; Page 143; 158pp; English.
XX CC ADB36346 to ADB36361 encode the human immune response associated proteins
XX CC given in ADB36330 to ADB36345, designated IRAP-1 to IRAP-16. IRAP
XX CC sequences have anti-HIV, anti-allergic, antianemic, antiasthmatic,
XX CC antiarteriosclerotic, dermatological, antiinflammatory, antidiabetic,
XX CC nephrotropic, antithyroid, thyromimetic, immunosuppressive, virucide,
XX CC antirheumatic, antiarthritic, osteopathic, antibacterial, virucide,
XX CC antiparasitic, protozoacide, fungicide, cerebroprotective,
XX CC neuroprotective, nootropic, antiparkinsonian, antipsoriatic, cytostatic
XX CC and cardiant activities, and can be used in gene therapy. The IRAP
XX CC proteins and polynucleotides can be used in diagnosing, preventing or
XX CC treating diseases or conditions associated with the decreased expression
XX CC or overexpression of IRAP, such as immune system, neurological,
XX CC developmental, muscle or cell proliferative disorders. The disorders may
XX CC include AIDS, allergies, anaemia, asthma, atherosclerosis, contact
XX CC dermatitis, Crohn's disease, diabetes, glomerulonephritis, Grave's
XX CC disease, Hashimoto's thyroiditis, multiple sclerosis, rheumatoid
XX CC arthritis, osteoporosis, systemic lupus erythematosus, infections (e.g.
XX CC bacterial, viral, fungal, parasitic or protozoal), stroke, Alzheimer's
XX CC disease, Parkinson's disease, psoriasis, cancer and cardiomyopathy. They
XX CC are also useful in assessing the effects of exogenous compounds on the
XX CC expression of nucleic acid and amino acid sequences of IRAP. The IRAP or
XX CC its fragments are useful in screening compounds for effectiveness as
XX CC agonist or antagonist of the polypeptides, or in altering the expression
XX CC of the target polynucleotide and compounds that specifically bind to or
XX CC modulate the activity of the polypeptide.

```

CC for diagnosing, treating and/or preventing cell proliferative disorders
 CC (e.g. actinic keratosis, arteriosclerosis, bursitis, cirrhosis, cancers
 CC or hepatitis), reproductive disorders (e.g. infertility, endometriosis,
 CC ovulatory defects, ectopic pregnancy or teratogenesis), gastritis, indigestion,
 CC disorders (e.g. dysphagia, peptide oesophagitis, gastritis, indigestion,
 CC anorexia or gastroenteritis), neurological disorders (e.g. Alzheimer's
 CC disease, Pick's disease, Huntington's disease or dementia), urologic
 CC disorders (e.g. infection, urethritis, cystitis, prostatitis or upper
 CC tract infections) and renal disorders (e.g. glomerulonephritis, renal
 CC amyloidosis, renal failure, Addison's disease or hypertension). The
 CC invention is also useful in gene therapy. The present sequence is human
 CC ORGA protein
 CC
 CC SQ Sequence 162 AA;
 Query Match 55.0%; Score 719.5; DB 6; Length 162;
 Best Local Similarity 89.1%; Pred. No. 1.2e-69;
 Matches 139; Conservative 0; Mismatches 6; Indels 11; Gaps 1;
 QY 1 MOTCPPLAFPGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSCNISNA 60
 DB 1 MOTCPPLAFPGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSCNISNA 60
 QY 61 FSHVNIKLRAHQESAI FNEVAPGYFSDGWLQVGGVAOLVIKGARDSHAGLYMHLV 120
 DB 61 FSHVNIKLRAHQESAI FNEVAPGYFSDGWLQVGGVAOLVIKGARDSHAGLYMHLV 120
 QY 121 GHORNNRQVTLVSEVLPPTPDEGRSPQSGSPAG 156
 DB 121 GHORNNRQVTLVSEVLPPTPDEGRSPQSGSPAG 156
 RESULT 9
 AAO29895
 ID AAO29895 standard; protein; 101 AA.
 AC AAO29895;
 XX
 DT 03-SEP-2003 (first entry)
 DE Human organelle-associated protein (ORGA) -8.
 KW Human; organelle-associated protein; ORGA; cell proliferative disorder;
 KW actinic keratosis; arteriosclerosis; reproductive disorder; infertility;
 KW endometriosis; gastrointestinal disorder; ectopic pregnancy; hepatitis;
 KW gastroenteritis; neurological disorder; Alzheimer's disease; cirrhosis;
 KW glomerulonephritis; renal amyloidosis; renal failure; Addison's disease;
 KW renal disorder; ovulatory defect; teratogenesis; Pick's disease; cancer;
 KW Huntington's disease; urologic disorder; cystitis; dementia; dysphagia;
 KW indigestion; gastritis; hypertension; gene therapy; bursitis; anorexia;
 KW infection; urethritis; prostatitis.
 OS Homo sapiens.
 XX
 PN WO2003044171-A2.
 XX
 PD 30-MAY-2003.
 XX
 XX 15-NOV-2002; 2002WO-US036807.
 XX
 XX 16-NOV-2001; 2001US-0332384P.
 PR 13-DEC-2001; 2001US-0341187P.
 PR 23-JAN-2002; 2002US-0351151P.
 PR 27-FEB-2002; 2002US-0360269P.
 PR 05-APR-2002; 2002US-0370637P.
 PR 14-JUN-2002; 2002US-0388946P.
 XX
 XX (INCY-) INCYTE GENOMICS INC.
 PA
 XX Ramkumar J, Burford N, Xu Y, Chawla NK, Kable AE, Becha SD;
 PI Richardson TW, Gorvad AE, Sprague WW, Elliott VS, Khare R;
 PI Hawkins PR, Jin P, Lee SY, Marquis JP, Swarnakar A, Wilson AD;
 PI Hafalia AJA, Bulloch S;

XX SQ Sequence 149 AA;
 Query Match 55.1%; Score 720.5; DB 7; Length 149;
 Best Local Similarity 97.2%; Pred. No. 8e-70;
 Matches 138; Conservative 1; Mismatches 2; Indels 1; Gaps 1;
 QY 1 MOTCPPLAFPGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSCNISNA 60
 DB 1 MOTCPPLAFPGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSCNISNA 60
 QY 61 FSHVNIKLRAHQESAI FNEVAPGYFSDGWLQVGGVAOLVIKGARDSHAGLYMHLV 120
 DB 61 FSHVNIKLRAHQESAI FNEVAPGYFSDGWLQVGGVAOLVIKGARDSHAGLYMHLV 120
 QY 121 GHORNNRQVTLVSEVLPPTPDEGRSPQSGSPAG 141
 DB 121 GHORNNRQVTLVSEVLPPTPDEGRSPQSGSPAG 142
 RESULT 8
 AAO29896
 ID AAO29896 standard; protein; 162 AA.
 AC AAO29896;
 XX
 DT 03-SEP-2003 (first entry)
 DE Human organelle-associated protein (ORGA) -9.
 KW Human; organelle-associated protein; ORGA; cell proliferative disorder;
 KW actinic keratosis; arteriosclerosis; reproductive disorder; infertility;
 KW endometriosis; gastrointestinal disorder; ectopic pregnancy; hepatitis;
 KW gastroenteritis; neurological disorder; Alzheimer's disease; cirrhosis;
 KW glomerulonephritis; renal amyloidosis; renal failure; Addison's disease;
 KW renal disorder; ovulatory defect; teratogenesis; Pick's disease; cancer;
 KW Huntington's disease; urologic disorder; cystitis; dementia; dysphagia;
 KW indigestion; gastritis; hypertension; gene therapy; bursitis; anorexia;
 KW infection; urethritis; prostatitis.
 OS Homo sapiens.
 XX
 PN WO2003044171-A2.
 XX
 PD 30-MAY-2003.
 XX
 XX 15-NOV-2002; 2002WO-US036807.
 XX
 XX 16-NOV-2001; 2001US-0332384P.
 PR 13-DEC-2001; 2001US-0341187P.
 PR 23-JAN-2002; 2002US-0351151P.
 PR 27-FEB-2002; 2002US-0360269P.
 PR 05-APR-2002; 2002US-0370637P.
 PR 14-JUN-2002; 2002US-0388946P.
 XX
 XX (INCY-) INCYTE GENOMICS INC.
 PA
 XX Ramkumar J, Burford N, Xu Y, Chawla NK, Kable AE, Becha SD;
 PI Richardson TW, Gorvad AE, Sprague WW, Elliott VS, Khare R;
 PI Hawkins PR, Jin P, Lee SY, Marquis JP, Swarnakar A, Wilson AD;
 PI Hafalia AJA, Bulloch S;
 XX
 XX WPI; 2003-457603/43.
 DR N-P8DB; AAL60349.
 XX
 XX New organelle-associated proteins and polynucleotides, useful for
 PT diagnosing, treating and/or preventing cell proliferative, reproductive,
 PT gastrointestinal, neurological, urologic, and renal disorders.
 XX
 XX Claim 1; Page 166; 194pp; English.
 PS
 XX The invention relates to human organelle-associated proteins (ORGA) and
 CC nucleic acid molecules encoding such proteins. ORGA sequences are useful

```

XX WPI; 2003-457603/43.
DR N-PSDB; AAL60548.
XX
PT New organelle-associated proteins and polynucleotides, useful for
PT diagnosing, treating and/or preventing cell proliferative, reproductive,
PT gastrointestinal, neurological, urologic, and renal disorders.
XX
PS Claim 1; Page 165-166; 194pp; English.
XX
CC The invention relates to human organelle-associated proteins (ORGA) and
CC nucleic acid molecules encoding such proteins. ORGA sequences are useful
CC for diagnosing, treating and/or preventing cell proliferative disorders
CC (e.g. actinic keratosis, arteriosclerosis, bursitis, cirrhosis, cancers
CC or hepatitis), reproductive disorders (e.g. infertility, endometriosis,
CC ovulatory defects, ectopic pregnancy or teratogenesis), gastrointestinal
CC disorders (e.g. dysphagia, peptide oesophagitis, gastritis, indigestion,
CC anorexia or gastroenteritis), neurological disorders (e.g. Alzheimer's
CC disease, Pick's disease, Huntington's disease or dementia), urologic
CC disorders (e.g. infection, urethritis, cystitis, prostatitis or upper
CC tract infections) and renal disorders (e.g. glomerulonephritis, renal
CC amyloidosis, renal failure, Addison's disease or hypertension). The
CC invention is also useful in gene therapy. The present sequence is human
CC ORGA protein
XX
SQ Sequence 101 AA;
    Query Match          37.6%; Score 492; DB 6; Length 101;
    Best Local Similarity 100.0%; Pred. No. 3.5e-45;
    Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
    QY 1 MOTCLPAPFGHVSQALGTLFLAASLSAQNEGWDSPICTEGVVSVWGENTVMSCNISNA 60
    Db 1 MOTCLPAPFGHVSQALGTLFLAASLSAQNEGWDSPICTEGVVSVWGENTVMSCNISNA 60
    QY 61 FSHVNIKLRAHQESAIENEVAPGYFSRDGMQ 92
    Db 61 FSHVNIKLRAHQESAIENEVAPGYFSRDGMQ 92
    RESULT 10
    AAB36660
    ID AAB36660 standard; protein; 212 AA.
    AC AAB36660;
    XX
    DT 13-MAR-2001 (first entry)
    DE Mouse K12 protein sequence SEQ ID NO:8.
    XX
    KW Mouse; CD7; K12; cognate ligand; cluster of differentiation; cancer;
    KW identification; inhibiting T cell proliferation; HIV; infection;
    KW activating natural killer cell proliferation; leukaemia; lymphoma;
    KW sepsis; graft versus host disease; autoimmune disease; arthritis;
    KW multiple sclerosis; rheumatoid arthritis; psoriatic arthritis; lupus;
    KW scleroderma; psoriasis; atopic dermatitis; type I diabetes mellitus;
    KW Hashimoto's thyroiditis; pernicious anaemia; Addison's disease; uveitis;
    KW myasthenia gravis; psoriasis; Guillain-Barre Syndrome; Grave's disease;
    KW systemic lupus erythematosus; dermatomyositis; asthma; eczema;
    KW atopic dermatitis; contact dermatitis; eczematous dermatitis;
    KW seborrheic dermatitis; rhinitis.
    XX
    OS Mus sp.
    XX
    PN WO200073333-A2.
    XX
    PD 07-DEC-2000.
    XX
    PF 26-MAY-2000; 2000WO-US014612.
    XX
    PR 28-MAY-1999; 99US-0136450P.
    XX
    PA (IMMUNEX CORP.

```

```

XX Lyman SD, Fanslow WC;
XX WPI; 2001-061511/07.
DR N-PSDB; AAC88154.
XX
PT Stimulating intracellular signaling of CD7 comprises contacting a cell
PT expressing CD7 with recombinant K12 protein, the cognate ligand of CD7,
PT to inhibit T cell proliferation and/or activate natural killer cell
PT proliferation.
XX
PS Claim 32; Page 42; 42pp; English.
XX
CC The present invention describes a method for stimulating (S) the
CC intracellular signalling of CD (cluster of differentiation) 7 comprising
CC contacting a cell that expresses CD7 with a recombinant K12 protein (1),
CC the cognate ligand of CD7. (S) is useful for inhibiting T cell
CC proliferation and/or activating NK (natural killer) cell proliferation of
CC and/or inducing NK toxicity in a mammal which involves administration of
CC K12 protein. It is also used for treating HIV-1 infection, cancer (T cell
CC leukaemia, acute lymphocytic leukaemia, cutaneous T cell lymphoma),
CC bacterial and viral infections, mediated by CD7. In the case of treating
CC T cell leukaemia the soluble K12 protein is covalently attached to a
CC toxin. A disease mediated by CD7 such as sepsis, graft versus host
CC disease due to transplantation, autoimmune diseases, multiple sclerosis,
CC arthritis, rheumatoid arthritis, psoriatic arthritis, scleroderma, lupus,
CC psoriasis, atopic dermatitis, type I diabetes mellitus, Hashimoto's
CC thyroiditis, pernicious anaemia, Addison's disease, myasthenia gravis,
CC lupus erythematosus and dermatomyositis, asthma, eczema, atopic
CC dermatitis, contact dermatitis, other eczematous dermatitides, seborrheic
CC dermatitis, and rhinitis is also treated by administering a K12
CC antagonist (neutralising antibody). The present sequence represents the
CC mouse K12 protein, which is given in the exemplification of the present
CC invention
XX
SQ Sequence 212 AA;
    Query Match          24.9%; Score 326; DB 4; Length 212;
    Best Local Similarity 37.8%; Pred. No. 1.2e-26;
    Matches 74; Conservative 25; Mismatches 73; Indels 24; Gaps 2;
    QY 10 GHVSQALGTLFLAASLSAQNEGWDSPICTEGVVSVWGENTVMSCNISNAFSHVNILKLR 69
    Db 10 GLFPRMLWALLLAASLSAQNHNVDPCCTEHEVSVNRGSRVVMACNLSNNLRDVTIELV 69
    QY 70 AHQGESAIENEVAPGYFSRDGMQLOVGGVAQLVKGARDSDHAGLYMHLVGHQNNRQV 129
    Db 70 TSEKTSIIFNHTPPGNYSKOSWQLHIQGVQVQLVITDAQDKHSGNYSWKLHGFQAEKFN 129
    QY 130 TLEVSGAEPQSPADTGFVPVP-----AVVTAFVILLVALVMFAWYRCRCQQ 176
    Db 130 NLTVNAADROKTEDELPTVKPDKPTAVRTEVILLIATITITIGVFWY-----182
    QY 177 RREKKFFLEPQMKVA 192
    Db 183 ----KQFPVPAQIQMS 194
    RESULT 11
    ADJ76167
    ID ADJ76167 standard; protein; 212 AA.
    AC ADJ76167;
    XX
    DT 20-MAY-2004 (first entry)
    DE Marker gene related amino acid sequence SEQ ID NO:1419.
    XX
    KW bronchial asthma; chronic obstructive pulmonary disease;
    KW respiratory epithelial cell; interleukin-13; respiratory; antiasthmatic;
    KW gene therapy; marker.
    XX

```

OS Mus musculus. 183 ---KQFPVAPQIQMS 194
XX ADJ76168
FN EPI394274-A2.
XX 03-MAR-2004.
PD 04-AUG-2003; 2003EP-00254857.
XX 06-AUG-2002; 2002JP-00229312.
XX 20-MAR-2003; 2003JP-00077212.
PR (GENO-) GENOX RES INC.
XX Ohtani N, Sugita Y, Yamaya M, Kubo H, Nagai H, Izuahara K;
XX WPI; 2004-193155/19.
XX Testing for bronchial asthma or chronic obstructive pulmonary disease by
XX comparing the expression level of a marker gene in a biological sample
XX from a subject with the expression level of the gene in a sample from a
XX healthy subject.
XX Claim 16; SEQ ID NO 1419; 241pp; English.
XX The present invention describes a method of testing for bronchial asthma
XX or chronic obstructive pulmonary disease. The method comprises
XX determining the expression level of a marker gene in a biological sample
XX from a subject, comparing the expression level determined with the
XX expression level of the marker gene in a biological sample from a healthy
XX subject, and judging whether the subject has bronchial asthma or chronic
XX obstructive pulmonary disease. The marker gene comprises: (a) a group of
XX genes (S1) whose expression levels increase when respiratory epithelial
XX cells are stimulated with interleukin-13; or (b) a group of genes (S2)
XX whose expression levels decrease when respiratory epithelial cells are
XX stimulated with interleukin-13. Also described: (1) a reagent (I) for
XX testing for bronchial asthma or chronic obstructive pulmonary disease;
XX (2) a kit for screening for a candidate compound for a therapeutic agent
XX to treat bronchial asthma or chronic obstructive pulmonary disease;
XX an animal model for bronchial asthma or chronic obstructive pulmonary
XX disease; (4) an inducer that induces bronchial asthma in a mouse; (5) a
XX method for producing an animal model for bronchial asthma or chronic
XX obstructive pulmonary disease; (6) a therapeutic agent for bronchial
XX asthma or chronic obstructive pulmonary disease, comprising the compound,
XX a marker gene or an antisense nucleic acid corresponding to a portion of
XX the marker gene, a ribozyme, a polynucleotide that suppresses the
XX expression of the gene through an RNAi effect or an antibody recognising
XX a protein encoded by a marker gene; and (7) a DNA chip for testing for
XX bronchial asthma or a chronic obstructive pulmonary disease, on which a
XX probe has been immobilised to assay a marker gene. (I) has respiratory
XX and antiasthmatic activities, and can be used in gene therapy. The method
XX is useful for testing for or screening for a therapeutic agent for
XX bronchial asthma or chronic obstructive pulmonary disease. The present
XX sequence is used in the exemplification of the present invention.

Query Match 24.9%; Score 326; DB 8; Length 212;
Best Local Similarity 37.8%; Pred. No. 1.2e-26;
Matches 74; Conservative 25; Mismatches 73; Indels 24; Gaps 2;
2;
QY 10 GHVSQALQTLFLAASLAQNEGSDSPICTEGVVSVSGNENTVMSCNISAFSHVNIKLR 69
DB 10 GLFPRMLWALLALLAASLNANDHNDWDEPCCTEHEVSUNRGRVVMACNISNNLRDVTIELV 69
QY 70 AHQGESAI FNEVAPGYFRDQWLOVQGGVAGLVITKGARDSHAGLYVHMLVGHORNNRQV 129
DB 70 TSEKTSIIENFTPPGNSKDSNQHLIQGVQQLVITDAQDGHGNSYKWLGLGFGFAEFNF 129
QY 130 TLEVSGAPQSPADPTGFWPVP-----AVVTAFVILVALVMPFAWYRCRCSQ 176
DB 130 NUTVNAQRQKTEDLPVTKVPDPPTAVRTVEVIIIIATITTIITGIGVFVWY----- 182
QY 177 REKKFFLLEPQMKVA 192

Db 183 ---KQFPVAPQIQMS 194
RESULT 12
ADJ76168
ID ADJ76168 standard; protein; 212 AA.
XX AC ADJ76168;
XX 20-MAY-2004 (first entry)
XX Marker gene related amino acid sequence SEQ ID NO:1420.
XX bronchial asthma; chronic obstructive pulmonary disease;
XX respiratory epithelial cell; interleukin-13; respiratory; antiasthmatic;
XX gene therapy; marker.
XX Mus musculus.
XX EPI394274-A2.
XX 03-MAR-2004.
XX 04-AUG-2003; 2003EP-00254857.
XX 06-AUG-2002; 2002JP-00229312.
XX 20-MAR-2003; 2003JP-00077212.
XX (GENO-) GENOX RES INC.
XX Ohtani N, Sugita Y, Yamaya M, Kubo H, Nagai H, Izuahara K;
XX WPI; 2004-193155/19.
XX Testing for bronchial asthma or chronic obstructive pulmonary disease by
XX comparing the expression level of a marker gene in a biological sample
XX from a subject with the expression level of the gene in a sample from a
XX healthy subject.
XX Claim 16; SEQ ID NO 1420; 241pp; English.
XX The present invention describes a method of testing for bronchial asthma
XX or chronic obstructive pulmonary disease. The method comprises
XX determining the expression level of a marker gene in a biological sample
XX from a subject, comparing the expression level determined with the
XX expression level of the marker gene in a biological sample from a healthy
XX subject, and judging whether the subject has bronchial asthma or chronic
XX obstructive pulmonary disease. The marker gene comprises: (a) a group of
XX genes (S1) whose expression levels increase when respiratory epithelial
XX cells are stimulated with interleukin-13; or (b) a group of genes (S2)
XX whose expression levels decrease when respiratory epithelial cells are
XX stimulated with interleukin-13. Also described: (1) a reagent (I) for
XX testing for bronchial asthma or chronic obstructive pulmonary disease;
XX (2) a kit for screening for a candidate compound for a therapeutic agent
XX to treat bronchial asthma or chronic obstructive pulmonary disease;
XX an animal model for bronchial asthma or chronic obstructive pulmonary
XX disease; (4) an inducer that induces bronchial asthma in a mouse; (5) a
XX method for producing an animal model for bronchial asthma or chronic
XX obstructive pulmonary disease; (6) a therapeutic agent for bronchial
XX asthma or chronic obstructive pulmonary disease, comprising the compound,
XX a marker gene or an antisense nucleic acid corresponding to a portion of
XX the marker gene, a ribozyme, a polynucleotide that suppresses the
XX expression of the gene through an RNAi effect or an antibody recognising
XX a protein encoded by a marker gene; and (7) a DNA chip for testing for
XX bronchial asthma or a chronic obstructive pulmonary disease, on which a
XX probe has been immobilised to assay a marker gene. (I) has respiratory
XX and antiasthmatic activities, and can be used in gene therapy. The method
XX is useful for testing for or screening for a therapeutic agent for
XX bronchial asthma or chronic obstructive pulmonary disease. The present
XX sequence is used in the exemplification of the present invention.

Query Match 24.8%; Score 325; DB 8; Length 212;
Best Local Similarity 37.8%; Pred. No. 1.5e-26;
Matches 74; Conservative 25; Mismatches 73; Indels 24; Gaps 2;
QY 10 GHVSQALGTLFLAASLSAQNKGWDSPTCTEGVSVSWGENTVMSCNISNAPSHVNIKL 69
DB 10 GLFPRLMALLLAASLAINVHVDKPCCTEHEVSVNRGSRVVMACNISNLRDVTIELV 69
QY 70 AHGQESALFNEVAPGVFSRDGWLQOVQGSVAOLVIKGRDASHAGLYMHLVCHQERNRQV 129
DB 70 TSKTISIFNQTPPNYSKDSWQLHIQGGQALVITDAQGRHSGBYWKLKGFQAEFRNF 129
QY 130 TLEVSGABPQSPADPTGFMPVP-----AVVTAVFILLVALVNFAMVYRCRCQQ 176
DB 130 NLIVNAADRQKTEDLPTVKPDKPTAVRTEVIIIAIATTTIITGIVGVWY----- 182
QY 177 RREKFFLEPQMKVA 192
DB 183 ----KQFFVAPQIQMS 194

RESULT 13
AAO29898
ID AAO29898 standard; protein; 107 AA.
XX
AC AAO29898;
XX
DT 03-SEP-2003 (first entry)
XX
DE Human organelle-associated protein (ORGA)-11.
XX
KW Human; organelle-associated protein; ORGA; cell proliferative disorder;
KW actinic keratosis; arteriosclerosis; reproductive disorder; infertility;
KW endometriosis; gastrointestinal disorder; ectopic pregnancy; hepatitis;
KW gastroenteritis; neurological disorder; Alzheimer's disease; cirrhosis;
KW glomerulonephritis; renal amyloidosis; renal failure; Addison's disease;
KW renal disorder; ovulatory defect; teratogenesis; Pick's disease; cancer;
KW Huntington's disease; urologic disorder; cystitis; dementia; dysphagia;
KW indigestion; gastritis; hypertension; gene therapy; bursitis; anorexia;
KW infection; urethritis; prostatitis.
XX
OS Homo sapiens.
XX
PN WO200304171-A2.
XX
PD 30-MAY-2003.
XX
PP 15-NOV-2002; 2002WO-US036807.
XX
PR 16-NOV-2001; 2001US-0332384P.
PR 13-DEC-2001; 2001US-0341187P.
PR 23-JAN-2002; 2002US-0351151P.
PR 27-FEB-2002; 2002US-0360269P.
PR 05-APR-2002; 2002US-0370637P.
PR 14-JUN-2002; 2002US-0388946P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Ramkumar J, Burford N, Xu Y, Chawla NK, Kable AE, Becha SD;
PI Richardson TW, Gorvad AE, Sprague WM, Elliott VS, Khare R;
PI Hawkins PJ, Jin P, Lee SY, Marquis JP, Swarnakar A, Wilson AD;
PI Hafalia AJA, Bulloch S;
XX
DR WPI; 2003-457603/43.
DR N-P5DB; AAL60551.
XX
PT New organelle-associated proteins and polynucleotides, useful for
PT diagnosing, treating and/or preventing cell proliferative, reproductive,
PT gastrointestinal, neurological, urologic, and renal disorders.
XX
PS Claim 1; Page 167; 194pp; English.

CC The invention relates to human organelle-associated proteins (ORGA) and
CC nucleic acid molecules encoding such proteins. ORGA sequences are useful
CC for diagnosing, treating and/or preventing cell proliferative disorders
CC (e.g. actinic keratosis, arteriosclerosis, bursitis, cirrhosis, cancers
CC or hepatitis), reproductive disorders (e.g. infertility, endometriosis,
CC ovulatory defects, ectopic pregnancy or teratogenesis), gastrointestinal
CC disorders (e.g. dysphagia, peptide oesophagitis, gastritis, indigestion,
CC anorexia or gastroenteritis), neurological disorders (e.g. Alzheimer's
CC disease, Pick's disease, Huntington's disease or dementia), urologic
CC disorders (e.g. infection, urethritis, cystitis, prostatitis or upper
CC tract infections) and renal disorders (e.g. glomerulonephritis, renal
CC amyloidosis, renal failure, Addison's disease or hypertension). The
CC invention is also useful in gene therapy. The present sequence is human
CC ORGA protein
XX
SQ Sequence 107 AA;

Query Match 22.5%; Score 294; DB 6; Length 107;
Best Local Similarity 100.0%; Pred. No. 1.4e-23;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MOTCPAPFGHVSOALGTLLFLAASLSAQNKGWDSPTCTEGVSVSWGENTVMS 55
DB 1 MOTCPAPFGHVSOALGTLLFLAASLSAQNKGWDSPTCTEGVSVSWGENTVMS 55

RESULT 14
AAB34671
ID AAB34671 standard; protein; 55 AA.
XX
AC AAB34671;
XX
DT 26-JAN-2001 (first entry)
XX
DE Gene 35 human secreted protein homologous amino acid sequence #155.
XX
KW Human; secreted protein; diagnosis; immunosuppressive; antiarthritic;
KW antirheumatic; antiproliferative; cytostatic; cardiac; vasotropic;
KW cerebroprotective; nootropic; neuroprotective; antibacterial; virucide;
KW fungicide; ophthalmological; gene therapy; autoimmune disease; neoplasm;
KW hyperproliferative disorder; cancer; cardiovascular disorder; infection;
KW cerebrovascular disorder; angiogenesis; nervous system disorder;
KW ocular disorder; wound healing; skin aging; food additive; preservative.
XX
OS Homo sapiens.
XX
PN WO2000056751-A1.
XX
PD 28-SEP-2000.
XX
PP 09-MAR-2000; 2000WO-US006013.
XX
PR 19-MAR-1999; 99US-0125360P.
PR 11-JUN-1999; 99US-0138626P.
PR 03-DEC-1999; 99US-0168662P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM, Komatsoulis G;
XX
DR WPI; 2000-579482/54.
XX
PT Isolated nucleic acid molecule encoding a human secreted protein is used
PT in preventing, treating or ameliorating a medical condition.
XX
PS Disclosure; Page 412; 419pp; English.
XX
CC The polynucleotide sequences given in AAC59738 to AAC59787 encode the
CC human secreted proteins given in AAB34577 to AAB34626. AAB34627 to
CC AAB34686 represent human secreted polypeptide sequences and proteins
CC homologous to them, which are given in the exemplification of the present
CC invention. Human secreted proteins have activities based on the tissues
CC and cells the genes are expressed in. Example of activities include:

antiarthritic; immunosuppressive; antirheumatic; antiproliferative; cyostatic; cardiant; vasotropic; cerebroprotective; nootropic; neuroprotective; antibacterial; virucide; fungicide; and ophthalmological. The polynucleotides and proteins can be used to prevent, treat or ameliorate a medical condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used in diagnosing a pathological condition or susceptibility to a pathological condition. Disorders which are diagnosed or treated include autoimmune diseases, hyperproliferative disorders e.g. neoplasms and cancers of the breast or liver, cardiovascular disorders, infections caused by disorders, angiogenesis, nervous system disorders. The proteins can also be used to aid wound healing and epithelial cell proliferation, to prevent skin aging due to sunburn, to maintain organs before transplantation, for supporting cell culture of primary tissues, to regenerate tissues and in chemotaxis. The proteins can also be used as food additive or preservative to increase or decrease storage capabilities. AAC59729 to AAC59737 and AAB34576 represent sequences used in the exemplification of the present invention

XX Sequence 55 AA;

Query Match 21.9%; Score 286; DB 3; Length 55;
Best Local Similarity 100.0%; Pred. No. 3.9e-23;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 194 LRAGAQQGLSRASAEIWTDPSEPTPRPLALVFKPSPLGALLELLSPQLFPYAADP 248
DB 1 LRAGAQQGLSRASAEIWTDPSEPTPRPLALVFKPSPLGALLELLSPQLFPYAADP 55

RESULT 15
AAB34672
ID AAB34672 standard; protein; 55 AA.
XX AAB34672;
AC AAB34672;
XX 26-JAN-2001 (first entry)
XX Human secreted protein sequence encoded by gene 35 SEQ ID NO:156.
XX Human: secreted protein; diagnosis; immunosuppressive; antirheumatic; antiproliferative; cyostatic; cardiant; vasotropic; cerebroprotective; nootropic; neuroprotective; antibacterial; virucide; fungicide; ophthalmological; gene therapy; autoimmune disease; neoplasm; hyperproliferative disorder; cancer; cardiovascular disorder; infection; cerebrovascular disorder; angiogenesis; nervous system disorder; ocular disorder; wound healing; skin aging; food additive; preservative.
XX Homo sapiens.
XX WO200056751-A1.
XX 28-SEP-2000.
XX 09-MAR-2000; 2000WO-US006013.
XX 19-MAR-1999; 99US-0125360P.
XX 11-JUN-1999; 99US-0138626P.
XX 03-DEC-1999; 99US-0168662P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Ruben SM, Komatsoulis G;
XX WPI; 2000-579482/54.
XX Isolated nucleic acid molecule encoding a human secreted protein is used in preventing, treating or ameliorating a medical condition.
XX Disclosure; Page 412-413; 419pp; English.
XX The polynucleotide sequences given in AAC59738 to AAC59787 encode the

CC human secreted proteins given in AAB34577 to AAB34626. AAB34627 to AAB34686 represent human secreted polypeptide sequences and proteins homologous to them, which are given in the exemplification of the present invention. Human secreted proteins have activities based on the tissues and cells the genes are expressed in. Example of activities include: antiarthritic; immunosuppressive; antirheumatic; antiproliferative; cyostatic; cardiant; vasotropic; cerebroprotective; nootropic; neuroprotective; antibacterial; virucide; fungicide; and ophthalmological. The polynucleotides and proteins can be used to prevent, treat or ameliorate a medical condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used in diagnosing a pathological condition or susceptibility to a pathological condition. Disorders which are diagnosed or treated include autoimmune diseases, hyperproliferative disorders e.g. neoplasms and cancers of the breast or liver, cardiovascular disorders, infections caused by disorders, angiogenesis, nervous system disorders. The proteins can also be used to aid wound healing and epithelial cell proliferation, to prevent skin aging due to sunburn, to maintain organs before transplantation, for supporting cell culture of primary tissues, to regenerate tissues and in chemotaxis. The proteins can also be used as food additive or preservative to increase or decrease storage capabilities. AAC59729 to AAC59737 and AAB34576 represent sequences used in the exemplification of the present invention

XX Sequence 55 AA;

Query Match 21.3%; Score 279; DB 3; Length 55;
Best Local Similarity 98.2%; Pred. No. 2.3e-22;
Matches 54; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 194 LRAGAQQGLSRASAEIWTDPSEPTPRPLALVFKPSPLGALLELLSPQLFPYAADP 248
DB 1 LRAGAQQGLSRASAEIWTDPSEPTPRPLALVFKPSPLGALLELLSPQLFPYAADP 55

Search completed: March 20, 2005, 09:47:33
Job time : 91 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 20, 2005, 09:25:08 ; Search time 27 seconds
(without alignments)
685.665 Million cell updates/sec

Title: US-10-080-522-1

Perfect score: 1308

Sequence: 1 MOTCLAPFGHVSQLGTLTLL.....PLGALELLSQPLFPFYAADP 248

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents, AA.*

1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*

2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*

3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*

4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*

5: /cgn2_6/ptodata/1/iaa/PCITUS_COMB.pep.*

6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1308	100.0	248	US-08-755-559-1	Sequence 1, Appli
2	1308	100.0	248	US-09-210-474-1	Sequence 1, Appli
3	1308	100.0	248	US-09-539-774-1	Sequence 1, Appli
4	1308	100.0	248	US-09-997-165-4	Sequence 4, Appli
5	1308	100.0	248	US-09-949-016-6297	Sequence 6297, Ap
6	1308	100.0	258	US-09-949-016-8807	Sequence 8807, Ap
7	326	24.9	212	US-09-997-165-8	Sequence 8, Appli
8	90.5	6.9	1171	US-09-489-039A-10298	Sequence 10298, A
9	84.5	6.5	381	US-09-252-991A-27771	Sequence 27771, A
10	84.5	6.5	455	US-09-949-016-6949	Sequence 6949, Ap
11	84.5	6.5	455	US-09-949-016-11026	Sequence 11026, A
12	83.5	6.4	199	US-09-430-503-46	Sequence 46, Appli
13	83.5	6.4	199	US-09-430-503-48	Sequence 48, Appli
14	83.5	6.4	1912	US-09-949-016-10490	Sequence 10490, A
15	83	6.3	152	US-09-270-767-43788	Sequence 43788, A
16	82.5	6.3	288	US-08-396-650-1	Sequence 1, Appli
17	82.5	6.3	288	US-08-768-626-1	Sequence 1, Appli
18	82.5	6.3	288	US-09-645-069-12	Sequence 12, Appli
19	81.5	6.2	199	US-09-430-503-42	Sequence 42, Appli
20	81.5	6.2	199	US-09-430-503-44	Sequence 44, Appli
21	81.5	6.2	209	US-09-430-503-18	Sequence 18, Appli
22	81.5	6.2	209	US-09-430-503-20	Sequence 20, Appli
23	81.5	6.2	269	US-09-430-503-2	Sequence 2, Appli
24	81.5	6.2	269	US-09-430-503-4	Sequence 4, Appli
25	81.5	6.2	269	US-09-430-503-6	Sequence 6, Appli
26	81.5	6.2	269	US-09-430-503-8	Sequence 8, Appli
27	81.5	6.2	310	US-09-328-352-5485	Sequence 5485, Ap

28	81.5	6.2	313	4	US-09-949-016-10974	Sequence 10974, A
29	81.5	6.2	370	4	US-09-252-991A-27810	Sequence 27810, A
30	81	6.2	583	4	US-09-328-352-6422	Sequence 6422, Ap
31	80.5	6.2	209	3	US-09-430-503-22	Sequence 22, Appli
32	80.5	6.2	209	3	US-09-430-503-24	Sequence 24, Appli
33	80.5	6.2	310	4	US-09-949-016-8501	Sequence 8501, Ap
34	80.5	6.2	310	4	US-09-949-016-8896	Sequence 8896, Ap
35	79.5	6.1	343	4	US-09-252-991A-22974	Sequence 22974, A
36	79	6.0	205	4	US-09-949-016-9980	Sequence 9980, Ap
37	79	6.0	205	4	US-09-949-016-9981	Sequence 9981, Ap
38	79	6.0	205	4	US-09-949-016-9982	Sequence 9982, Ap
39	79	6.0	485	4	US-09-902-540-11019	Sequence 11019, A
40	79	6.0	545	4	US-10-237-551-121	Sequence 121, App
41	79	6.0	545	4	US-10-237-551-157	Sequence 157, App
42	79	6.0	547	4	US-10-237-551-216	Sequence 216, App
43	79	6.0	618	4	US-09-970-516-4	Sequence 4, Appli
44	79	6.0	618	4	US-09-817-676A-14	Sequence 14, Appli
45	78.5	6.0	225	4	US-09-530-139-42	Sequence 42, Appli

ALIGNMENTS

RESULT 1
US-08-755-559-1
; Sequence 1, Application US/08755559
; Patent No. 5912142
; GENERAL INFORMATION:
; APPLICANT: KAUFMAN, RUSSEL E.
; TITLE OF INVENTION: GENE PRODUCT OVER EXPRESSED IN CANCER
; TITLE OF INVENTION: CELLS
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHUYE P.C.
; STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/755,559
; FILING DATE: 22-NOV-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: WILSON, MARY J.
; REGISTRATION NUMBER: 32,955
; REFERENCE/DOCKET NUMBER: 1579-116
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 248 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-755-559-1

Query Match 100.0%; Score 1308; DB 2; Length 248;
Best Local Similarity 100.0%; Pred. No. 2.6e-139;
Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MOTCLAPFGHVSQLGTLTLLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSCNINA 60
Db 1 MOTCLAPFGHVSQLGTLTLLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSCNINA 60

us-10-080-522-1-rai

Sun Mar 20 14:17:42 2005

QY 61 FSHVNIKLRAHQESAI FNEVAPGYFSRDGWLQVQGGVAQLVIKGARDSHAGLYMHLV 120
 DB 61 FSHVNIKLRAHQESAI FNEVAPGYFSRDGWLQVQGGVAQLVIKGARDSHAGLYMHLV 120
 QY 121 GHQRRNRQVTLVSGAEPQSDPTGFWPVPVAVTAVFILLVALVMPFAYRCRCQORREK 180
 DB 121 GHQRRNRQVTLVSGAEPQSDPTGFWPVPVAVTAVFILLVALVMPFAYRCRCQORREK 180
 QY 181 KFFLLEPQMKVAALRAGAQOGLSRASAEIWTDPSEPTPRPLALVPKPSPLGALLELLSPQ 240
 DB 181 KFFLLEPQMKVAALRAGAQOGLSRASAEIWTDPSEPTPRPLALVPKPSPLGALLELLSPQ 240
 QY 241 LFPYAADP 248
 DB 241 LFPYAADP 248

RESULT 2
 US-09-210-474-1
 ; Sequence 1, Application US/09210474
 ; Patent No. 6072034
 ; GENERAL INFORMATION:
 ; APPLICANT: KAUFMAN, RUSSEL E.
 ; APPLICANT: SLENTZ-KESLER, KIMBERLY
 ; TITLE OF INVENTION: GENE PRODUCT OVER EXPRESSED IN CANCER
 ; TITLE OF INVENTION: CELLS
 ; NUMBER OF SEQUENCES: 2
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: NIXON & VANDERHVE P.C.
 ; STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
 ; CITY: ARLINGTON
 ; STATE: VIRGINIA
 ; COUNTRY: U.S.A.
 ; ZIP: 22201-4714
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/210,474
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/755,559
 ; FILING DATE: 22-NOV-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: WILSON, MARY J.
 ; REGISTRATION NUMBER: 32,955
 ; REFERENCE/DOCKET NUMBER: 1579-116
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (703) 816-4000
 ; TELEFAX: (703) 816-4100
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 248 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS:
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-09-210-474-1

Query Match 100.0%; Score 1308; DB 3; Length 248;
 Best Local Similarity 100.0%; Pred. No. 2.6e-139;
 Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MOTCPALFPGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSCNISNA 60
 DB 1 MOTCPALFPGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSCNISNA 60
 QY 61 FSHVNIKLRAHQESAI FNEVAPGYFSRDGWLQVQGGVAQLVIKGARDSHAGLYMHLV 120
 DB 61 FSHVNIKLRAHQESAI FNEVAPGYFSRDGWLQVQGGVAQLVIKGARDSHAGLYMHLV 120

QY 121 GHQRRNRQVTLVSGAEPQSDPTGFWPVPVAVTAVFILLVALVMPFAYRCRCQORREK 180
 DB 121 GHQRRNRQVTLVSGAEPQSDPTGFWPVPVAVTAVFILLVALVMPFAYRCRCQORREK 180
 QY 181 KFFLLEPQMKVAALRAGAQOGLSRASAEIWTDPSEPTPRPLALVPKPSPLGALLELLSPQ 240
 DB 181 KFFLLEPQMKVAALRAGAQOGLSRASAEIWTDPSEPTPRPLALVPKPSPLGALLELLSPQ 240
 QY 241 LFPYAADP 248
 DB 241 LFPYAADP 248

RESULT 3
 US-09-539-774-1
 ; Sequence 1, Application US/09539774
 ; Patent No. 6350615
 ; GENERAL INFORMATION:
 ; APPLICANT: KAUFMAN, RUSSEL E.
 ; APPLICANT: SLENTZ-KESLER, KIMBERLY
 ; TITLE OF INVENTION: GENE PRODUCT OVER EXPRESSED IN CANCER
 ; TITLE OF INVENTION: CELLS
 ; NUMBER OF SEQUENCES: 2
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: NIXON & VANDERHVE P.C.
 ; STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
 ; CITY: ARLINGTON
 ; STATE: VIRGINIA
 ; COUNTRY: U.S.A.
 ; ZIP: 22201-4714
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/539,774
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 09/210,474
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: WILSON, MARY J.
 ; REGISTRATION NUMBER: 32,955
 ; REFERENCE/DOCKET NUMBER: 1579-116
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (703) 816-4000
 ; TELEFAX: (703) 816-4100
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 248 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS:
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-09-539-774-1

Query Match 100.0%; Score 1308; DB 3; Length 248;
 Best Local Similarity 100.0%; Pred. No. 2.6e-139;
 Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MOTCPALFPGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSCNISNA 60
 DB 1 MOTCPALFPGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSCNISNA 60
 QY 61 FSHVNIKLRAHQESAI FNEVAPGYFSRDGWLQVQGGVAQLVIKGARDSHAGLYMHLV 120
 DB 61 FSHVNIKLRAHQESAI FNEVAPGYFSRDGWLQVQGGVAQLVIKGARDSHAGLYMHLV 120
 QY 121 GHQRRNRQVTLVSGAEPQSDPTGFWPVPVAVTAVFILLVALVMPFAYRCRCQORREK 180
 DB 121 GHQRRNRQVTLVSGAEPQSDPTGFWPVPVAVTAVFILLVALVMPFAYRCRCQORREK 180

Db 121 GHORNRQVTLVSGAEPQSDPTGFWPVPVAVVTAFFILLVALVMFAWYRCRCQQORREK 180
Qy 181 KFFLLEPOMKVAALRAGAQOGLSRASAEIWTDPSETPRPLALVFKPSPGLGALELLSPQ 240
Db 181 KFFLLEPOMKVAALRAGAQOGLSRASAEIWTDPSETPRPLALVFKPSPGLGALELLSPQ 240
Qy 241 LFPYAADP 248
Db 241 LFPYAADP 248

RESULT 4
US-09-997-165-4
; Sequence 4, Application US/09997165
; Patent No. 6762030
; GENERAL INFORMATION:
; APPLICANT: Lyman, Stewart D.
; APPLICANT: Fanslow, William C.
; TITLE OF INVENTION: LIGAND FOR CD7 AND METHODS OF USE THEREOF
; FILE REFERENCE: 2913-US
; CURRENT APPLICATION NUMBER: US/09/997,165
; CURRENT FILING DATE: 2001-11-27
; PRIOR APPLICATION NUMBER: PCT/US00/14612
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/136,450
; PRIOR FILING DATE: 1999-05-28
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 248
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-997-165-4

Query Match 100.0%; Score 1308; DB 4; Length 248;
Best Local Similarity 100.0%; Pred. No. 2.6e-139;
Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MOTCLAPFGHVSQLGTLLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSNCISNA 60
Db 1 MOTCLAPFGHVSQLGTLLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSNCISNA 60
Qy 61 FSHVNIKLRAHQESAIENEVAPGYFSRDGWLQVQGGVAQLVIKIGARDSHAGLYMHLV 120
Db 61 FSHVNIKLRAHQESAIENEVAPGYFSRDGWLQVQGGVAQLVIKIGARDSHAGLYMHLV 120
Qy 121 GHORNRQVTLVSGAEPQSDPTGFWPVPVAVVTAFFILLVALVMFAWYRCRCQQORREK 180
Db 121 GHORNRQVTLVSGAEPQSDPTGFWPVPVAVVTAFFILLVALVMFAWYRCRCQQORREK 180
Qy 181 KFFLLEPOMKVAALRAGAQOGLSRASAEIWTDPSETPRPLALVFKPSPGLGALELLSPQ 240
Db 181 KFFLLEPOMKVAALRAGAQOGLSRASAEIWTDPSETPRPLALVFKPSPGLGALELLSPQ 240
Qy 241 LFPYAADP 248
Db 241 LFPYAADP 248

RESULT 5
US-09-949-016-6297
; Sequence 6297, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6297
; LENGTH: 248
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6297

Query Match 100.0%; Score 1308; DB 4; Length 248;
Best Local Similarity 100.0%; Pred. No. 2.6e-139;
Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MOTCLAPFGHVSQLGTLLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSNCISNA 60
Db 1 MOTCLAPFGHVSQLGTLLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSNCISNA 60
Qy 61 FSHVNIKLRAHQESAIENEVAPGYFSRDGWLQVQGGVAQLVIKIGARDSHAGLYMHLV 120
Db 61 FSHVNIKLRAHQESAIENEVAPGYFSRDGWLQVQGGVAQLVIKIGARDSHAGLYMHLV 120
Qy 121 GHORNRQVTLVSGAEPQSDPTGFWPVPVAVVTAFFILLVALVMFAWYRCRCQQORREK 180
Db 121 GHORNRQVTLVSGAEPQSDPTGFWPVPVAVVTAFFILLVALVMFAWYRCRCQQORREK 180
Qy 181 KFFLLEPOMKVAALRAGAQOGLSRASAEIWTDPSETPRPLALVFKPSPGLGALELLSPQ 240
Db 181 KFFLLEPOMKVAALRAGAQOGLSRASAEIWTDPSETPRPLALVFKPSPGLGALELLSPQ 240
Qy 241 LFPYAADP 248
Db 241 LFPYAADP 248

RESULT 6
US-09-949-016-8807
; Sequence 8807, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8807
; LENGTH: 258
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-8807

Query Match 100.0%; Score 1308; DB 4; Length 258;
Best Local Similarity 100.0%; Pred. No. 2.7e-139;
Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MOTCLAPFGHVSQLGTLLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSNCISNA 60
Db 11 MOTCLAPFGHVSQLGTLLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSNCISNA 70
Qy 61 FSHVNIKLRAHQESAIENEVAPGYFSRDGWLQVQGGVAQLVIKIGARDSHAGLYMHLV 120
Db 71 FSHVNIKLRAHQESAIENEVAPGYFSRDGWLQVQGGVAQLVIKIGARDSHAGLYMHLV 130

QY 121 GHORNRQVTLVSGABPQSDPTGFWPVPVAVTAVFILLVALVFMFAYRCRCQQRRK 180
DB 131 GHORNRQVTLVSGABPQSDPTGFWPVPVAVTAVFILLVALVFMFAYRCRCQQRRK 190
QY 181 KFFLEPQMKVAALRAGAQOGLSRASAEIWTDPSEPTPRPLALVFKPSPIGALELLSPQ 240
DB 191 KFFLEPQMKVAALRAGAQOGLSRASAEIWTDPSEPTPRPLALVFKPSPIGALELLSPQ 250
QY 241 LFPYAADP 248
DB 251 LFPYAADP 258

RESULT 7
US-09-997-165-8
; Sequence 8, Application US/09997165
; Patent No. 6762030
; GENERAL INFORMATION:
; APPLICANT: Lyman, Stewart D.
; TITLE OF INVENTION: LIGAND FOR CD7 AND METHODS OF USE THEREOF
; FILE REFERENCE: 2913-US
; CURRENT APPLICATION NUMBER: US/09/997,165
; CURRENT FILING DATE: 2001-11-27
; PRIOR APPLICATION NUMBER: PCT/US00/14612
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/136,450
; PRIOR FILING DATE: 1999-05-28
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 8
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Mus sp.
US-09-997-165-8
Query Match 24.9%; Score 326; DB 4; Length 212;
Best Local Similarity 37.8%; Pred. No. 1.9e-28;
Matches 74; Conservative 25; Mismatches 73; Indels 24; Gaps 2;
QY 10 GHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSCNISNAFASHVNIKUR 69
DB 10 GLFPRMLWALLLSALNAHNDVWDEPCCTEHEVSVNRGSRVVMACNISNLRDVTIELV 69
QY 70 AHQGESAIENFVAPGYSRQWQLOVQGGVAQLVKGARDSHAGLYMHLVGHQRNRQV 129
DB 70 TSEKTSIIFNHTPPGNYSKDSWQLHIGVQVQALVITDAQDKHSGNYSWKLHGFQAEKFN 129
QY 130 TLEVSGAEPOSADPTGFWPVPV-----AVVTAVFILLVALVFMFAYRCRCQQ 176
DB 130 NLTVNAADRKTELDLPVKVDPKPTAVRTEVIIIAIATITIIITGIGVFMVY----- 182
QY 177 RREKKFFLEPQMKVA 192
DB 183 ----KQPPVAPQIQMS 194
RESULT 8
US-09-489-039A-10298
; Sequence 10298, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 10298
; LENGTH: 1171

; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-10298
Query Match 6.9%; Score 90.5; DB 4; Length 1171;
Best Local Similarity 22.7%; Pred. No. 1.1;
Matches 57; Conservative 37; Mismatches 96; Indels 61; Gaps 13;
QY 19 LFLAASLSAQNEGWDSPICTEGVSV-----VSWGENTV-----MSCNISNAFASHV 64
DB 582 LAFYAASL-ASHPQWRPLPV-DDGLVSVQVTRLIQRLGQNSSESTLYQKLAQVANOYADM 639
QY 65 NI-KLRAHQESAIENFVAPGYSRQWQLOVQGGVAQLVKGARDSHAGLYMHLV 121
DB 640 RIADMTADTDASRLFTSTDEVVFGMFTQAWEQAVQPAIEKVVAAE-RRDE-----MDMWLSD 694
QY 122 HORNRQVT-----LEVSGAEPOSADPTGFWPVPVAVTAVFILLVALVFMFAYRCRCQQ 177
DB 695 TKQTAAQSTSPALRAERLAEVYADFSGAW-----LDFLNSLRWQRAATLSDA 742
QY 178 REKKFFLEPQMK-VAALRAGAQOQ-----LSRASAEIWTDPSE- 216
DB 743 IDQLTLMADVRSQPLVALMNTLSVQRTGTGTGEAIADSLVKSARQLFNRDNPVIDQRSG 802
QY 217 TPRPLALVFKP 227
DB 803 ARGFLDATFGP 813
RESULT 9
US-09-252-991A-27771
; Sequence 27771, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27771
; LENGTH: 381
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27771
Query Match 6.5%; Score 84.5; DB 4; Length 381;
Best Local Similarity 26.2%; Pred. No. 0.91;
Matches 60; Conservative 27; Mismatches 91; Indels 51; Gaps 12;
QY 1 MOTCPALFPGHVSOALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSCNISNA 60
DB 192 MVTAPLVMPVEITGLSLLLFVA---MAQLGWQP---ERGIVTI-WIAHT-----SFC 238
QY 61 FSHVNIKLRAHQESAIENFVAPGYSRQWQ-----LQVQGGVAQLVKGARDSHAGLYM 117
DB 239 SSVVAVVVSARLRELDLSIEEAAMDGLGAPKWKVFLITIPMIAPSLAAGMMWFA-LSLD 297
QY 118 HLHGQRNRQVTLVSG-----AEQSPADPTGFWPVPVAVTAVFILLVAL-VMFAY 169
DB 298 DLV-----LASFVSGPGSTLMEVFSVRLGVKPEINAVASLILLSVSLFTFFAWY 349
QY 170 RCRCSQORREKKFFLEPQMKVAALRAGAQOGLSRASAEIWTDPSEPTP 218
DB 350 FTRQAEERR-----RRAIQAMEETATD-WOKGS-PTP 380
RESULT 10

Sun Mar 20 14:17:42 2005

Query Match 6.4%; Score 83.5; DB 3; Length 199;
Best Local Similarity 21.9%; Pred. No. 0.44;
Matches 34; Conservative 28; Mismatches 54; Indels 39; Gaps 6;
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 43788
LENGTH: 152
TYPE: PRT
ORGANISM: Drosophila melanogaster
FEATURE:
OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-43788

Query Match 6.3%; Score 83; DB 4; Length 152;
Best Local Similarity 33.3%; Pred. No. 0.33;
Matches 31; Conservative 9; Mismatches 27; Indels 26; Gaps 7;
QY 170 RCRCSQQRREKKFFLL---EPQMK-----VAALRAGAAQGLSRASAEIWTDPDSEPTP- 218
DB 51 RIMCAPPREAVHRIIRDAHPQMGDNRRPRRLAPLQIGNQ--SR-----KQPIPP 99
QY 219 --RP-LALVFKPSPLGALELLSPQLFPYAADP 248
DB 100 RRRPALALILAPSPFGSQ-IAPSPLLPVAPSP 131

Search completed: March 20, 2005, 09:50:12
Job time : 29 secs

Query Match 6.4%; Score 83.5; DB 3; Length 199;
Best Local Similarity 21.9%; Pred. No. 0.44;
Matches 34; Conservative 28; Mismatches 54; Indels 39; Gaps 6;
QY 39 TEGVVSVSW-----GENTVMSCNISNAPSHVNIKLRAHQESALFNEVAPGYFSDGWLQ 93
DB 66 TGLTSVSWSFQEGADTTVS-----FFHYS-----QGQVYIGNYPPFKDRISWAG 111
QY 94 QVQGVVAOLVIKARDSHAGLYMHL-----VGHORNROVTLVSGAEPQSA PDT 144
DB 112 DLDKDDASININMOFIHNGTYICDVKNPPDIVVQPGH-----IRIYVEKENLPVF 163
QY 145 GFVPVAVVTAVFI---LLVALVMFAWYRCRCQQ 176
DB 164 PWWVVGVITAVVLGTLISLILAVLYRRKNSKR 198

RESULT 14
US-09-949-016-10490
Sequence 10490, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10490
LENGTH: 1912
TYPE: PRT
ORGANISM: Human
US-09-949-016-10490

Query Match 6.4%; Score 83.5; DB 4; Length 1912;
Best Local Similarity 20.7%; Pred. No. 14;
Matches 62; Conservative 31; Mismatches 90; Indels 117; Gaps 15;
QY 5 PLAPPG--HVSQALGTLFLAASLSAONEGW-----DSPICTEGVSVSW-- 47
DB 592 PLNPPGTDHITVPL-ALMFEDVTVAATNFSFYDCSAVQALEAAAPCRACVGSIRCHWCP 650
QY 48 -----GENTVMSCNISNAPSHVNIKLRAHQESALFNEVAPGYFSDGWLQ 94
DB 651 QSSHCYVGEHCPEGERTIYSA-----QEVDIQVRGACPCQVEGLAGP-HLVPGWESH 703
QY 95 VQGVVAOLVIKARDSHAGLYMHLVGHORNROVTLVSGAEPQSA PDTG 146
DB 704 LALVRNL--OHFRGLPASHFCHWLELPGELGLPATLE-----ETAGDSGLIHCOAHQF 755
QY 147 -----WPVPVAVT-----AVFILLVALVMFAWYRCR-----CSQORREKKF 182
DB 756 YPMSQRELFPVIYVTOGEAORLDNTHALYVIL-----YDCAMGHDPDCH----- 800
QY 183 FLLEPQMKVAALRAGAAQGLSRASAEIWTDPDSEPTP--PLALVFKPSPLGALELLSPQ 240
DB 801 -----COANRSLGCLWCADGQACRYGLIC-----PPGAVELLCPAP 838

RESULT 15
US-09-270-767-43788
Sequence 43788, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 20, 2005, 09:49:19 ; Search time 74 Seconds
(without alignments)
1107.760 Million cell updates/sec

Title: US-10-080-522-1
Perfect score: 1308
Sequence: 1 MQTCLAPGHVSQALGTL.....PLGALELLSQPLFPFYAADP 248

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1401741 seqs, 330541175 residues

Total number of hits satisfying chosen parameters: 1401741

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 10%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1308	100.0	248	9 US-09-799-777-32	Sequence 32, Appl
2	1308	100.0	248	9 US-09-997-165-4	Sequence 4, Appl
3	1308	100.0	248	14 US-10-080-522-1	Sequence 1, Appl
4	326	24.9	212	9 US-09-997-165-8	Sequence 8, Appl
5	98	7.5	969	15 US-10-052-648A-35	Sequence 35, Appl
6	96	7.3	969	15 US-10-052-648A-34	Sequence 34, Appl
7	95.5	7.3	506	15 US-10-072-012-446	Sequence 446, App
8	91.5	7.0	597	15 US-10-425-114-62807	Sequence 62807, A
9	91	7.0	6620	15 US-10-080-334-290	Sequence 290, App
10	91	7.0	6620	15 US-10-408-765A-2291	Sequence 2291, Ap
11	91	7.0	7968	13 US-10-077-130-5	Sequence 5, Appl
12	89	6.8	435	10 US-09-759-130B-275	Sequence 275, App
13	89	6.8	435	14 US-10-189-123-5	Sequence 5, Appl

Sequence 5, Appl
Sequence 275, App
Sequence 273, App
Sequence 3, Appl
Sequence 3, Appl
Sequence 273, App
Sequence 14023, A
Sequence 43635, A
Sequence 24, Appl
Sequence 46, Appl
Sequence 48, Appl
Sequence 153, App
Sequence 678, App
Sequence 19, Appl
Sequence 13729, A
Sequence 107, App
Sequence 12, Appl
Sequence 12, Appl
Sequence 48, Appl
Sequence 135, App
Sequence 48, Appl
Sequence 59, Appl
Sequence 60, Appl
Sequence 218, App
Sequence 244, App
Sequence 4, Appl
Sequence 4, Appl
Sequence 14, Appl
Sequence 16, Appl
Sequence 18, Appl
Sequence 14, Appl
Sequence 16, Appl

US-10-188-495-5
US-10-741-790-275
US-09-759-130B-273
US-10-189-123-3
US-10-188-495-3
US-10-741-790-273
US-10-156-761-14023
US-10-282-122A-43635
US-10-322-696-24
US-10-095-131A-46
US-10-095-131A-48
US-10-091-438-153
US-09-764-853-678
US-10-626-832-19
US-10-369-493-13729
US-10-367-094-107
US-10-068-215-12
US-10-115-615-12
US-10-190-115-48
US-10-190-115-135
US-10-369-072-48
US-10-042-865-59
US-10-042-865-60
US-09-989-920-218
US-09-989-890-244
US-10-042-865-4
US-10-149-819-4
US-10-190-115-14
US-10-190-115-16
US-10-190-115-18
US-10-369-072-14
US-10-369-072-16

ALIGNMENTS

RESULT 1
US-09-799-777-32
; Sequence 32, Application US/09799777
; Patent No. US20020091244A1
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; Hillman, Jennifer L.
; Corley, Neil C.
; Guegler, Karl J.
; Baugh, Mariah
; Sather, Susan
; Shah, Purvi
; TITLE OF INVENTION: HUMAN SIGNAL PEPTIDE-CONTAINING PROTEINS
; NUMBER OF SEQUENCES: 154
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/799,777
; FILING DATE: 06-Mar-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/002,485
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: BILLINGS, LUCY J.
; REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0459 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 248 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 32 :
US-09-799-777-32

Query Match 100.0%; Score 1308; DB 9; Length 248;
Best Local Similarity 100.0%; Pred. No. 1.4e-124;
Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MOTCPAPFGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSNCISNA 60
Db 1 MOTCPAPFGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSNCISNA 60
QY 61 FSHVNIKLRAHQESAI FNEVAPGYFSRDGWLQVGGVAQLVIKARDSHAGLYMHV 120
Db 61 FSHVNIKLRAHQESAI FNEVAPGYFSRDGWLQVGGVAQLVIKARDSHAGLYMHV 120
QY 121 GHORNNRQVTLVSGAEPQSDPTGFWPVPVAVTAVFILLVALVMPFYRCRCQORREK 180
Db 121 GHORNNRQVTLVSGAEPQSDPTGFWPVPVAVTAVFILLVALVMPFYRCRCQORREK 180
QY 181 KFFLLEPQMKVAALRAGAQOGLSRASAEIWTDPSETPRPLALVPKPSPLGALLELLSQP 240
Db 181 KFFLLEPQMKVAALRAGAQOGLSRASAEIWTDPSETPRPLALVPKPSPLGALLELLSQP 240
QY 241 LFPYAADP 248
Db 241 LFPYAADP 248

RESULT 2
US-09-997-165-4
Sequence 4, Application US/09997165
Patent No. US20020141999A1
GENERAL INFORMATION:
APPLICANT: Lyman, Stewart D.
APPLICANT: Fanslow, William C.
TITLE OF INVENTION: LIGAND FOR CD7 AND METHODS OF USE THEREOF
FILE REFERENCE: 2913-US
CURRENT APPLICATION NUMBER: US/09/997,165
CURRENT FILING DATE: 2001-11-27
PRIOR APPLICATION NUMBER: PCT/US00/14612
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/136,450
PRIOR FILING DATE: 1999-05-28
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 4
LENGTH: 248
TYPE: PRT
ORGANISM: Homo sapiens
US-09-997-165-4

Query Match 100.0%; Score 1308; DB 9; Length 248;
Best Local Similarity 100.0%; Pred. No. 1.4e-124;
Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MOTCPAPFGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSNCISNA 60
Db 1 MOTCPAPFGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSNCISNA 60
QY 61 FSHVNIKLRAHQESAI FNEVAPGYFSRDGWLQVGGVAQLVIKARDSHAGLYMHV 120
Db 61 FSHVNIKLRAHQESAI FNEVAPGYFSRDGWLQVGGVAQLVIKARDSHAGLYMHV 120

QY 121 GHORNNRQVTLVSGAEPQSDPTGFWPVPVAVTAVFILLVALVMPFYRCRCQORREK 180
Db 121 GHORNNRQVTLVSGAEPQSDPTGFWPVPVAVTAVFILLVALVMPFYRCRCQORREK 180
QY 181 KFFLLEPQMKVAALRAGAQOGLSRASAEIWTDPSETPRPLALVPKPSPLGALLELLSQP 240
Db 181 KFFLLEPQMKVAALRAGAQOGLSRASAEIWTDPSETPRPLALVPKPSPLGALLELLSQP 240
QY 241 LFPYAADP 248
Db 241 LFPYAADP 248

RESULT 3
US-10-080-522-1
Sequence 1, Application US/10080522
Publication No. US20030096326A1
GENERAL INFORMATION:
APPLICANT: KAUFMAN, RUSSEL E.
APPLICANT: SLENTZ-KESLER, KIMBERLY
TITLE OF INVENTION: GENE PRODUCT OVER EXPRESSED IN CANCER CELLS
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHYE P.C.
STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/080,522
FILING DATE: 25-Feb-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/539,774
FILING DATE: 31-MAR-2000
APPLICATION NUMBER: US 09/210,474
FILING DATE: 14-DEC-1998
APPLICATION NUMBER: US 08/755,559
FILING DATE: 22-NOV-1996
ATTORNEY/AGENT INFORMATION:
NAME: WILSON, MARY J.
REGISTRATION NUMBER: 32,955
REFERENCE/DOCKET NUMBER: 1579-645
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 248 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-080-522-1

Query Match 100.0%; Score 1308; DB 14; Length 248;
Best Local Similarity 100.0%; Pred. No. 1.4e-124;
Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MOTCPAPFGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSNCISNA 60
Db 1 MOTCPAPFGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSNCISNA 60
QY 61 FSHVNIKLRAHQESAI FNEVAPGYFSRDGWLQVGGVAQLVIKARDSHAGLYMHV 120

Db 61 FSHVNIKURAGORSALFNEVAPGYFSRDGQVQGVQVLAQVTKGARDSHAGLYMMHLV 120
Qy 121 GHORNRQVTLVSGAEPQSDPTGFWFVPAVTVAVFILLVALVWFAYRCRCQQRRK 180
Db 121 GHORNRQVTLVSGAEPQSDPTGFWFVPAVTVAVFILLVALVWFAYRCRCQQRRK 180
Qy 181 KFFLEPQMKVAALRAGAQOGLSRASAEIWTPTDSEPTPRPLALVFKPSPLGALLELSPQ 240
Db 181 KFFLEPQMKVAALRAGAQOGLSRASAEIWTPTDSEPTPRPLALVFKPSPLGALLELSPQ 240
Qy 241 LFPYAADP 248
Db 241 LFPYAADP 248

RESULT 4

US-09-997-165-8
; Sequence 8, Application US/09997165
; Patent No. US20020141999A1
; GENERAL INFORMATION:
; APPLICANT: Lyman, Stewart D.
; APPLICANT: Farnlow, William C.
; TITLE OF INVENTION: LGAND FOR CD7 AND METHODS OF USE THEREOF
; FILE REFERENCE: 2913-US
; CURRENT APPLICATION NUMBER: US/09/997,165
; CURRENT FILING DATE: 2001-11-27
; PRIOR APPLICATION NUMBER: PCT/US00/14612
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/136,450
; PRIOR FILING DATE: 1999-05-28
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 8
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Mus sp.
; US-09-997-165-8

Query Match 24.9%; Score 326; DB 9; Length 212;
Best Local Similarity 37.8%; Pred. No. 1.1e-24;
Matches 74; Conservative 25; Mismatches 73; Indels 24; Gaps 2;
Qy 10 GHVSQALGTLFLAASLAQNEGWDSPTCTGVSWSGENTVMSNISAFSHVNIKL 69
Db 10 GLFPRMLWALLLAASLAHNDVNDPEPCCTHEVSVNRGRVWMACTNLRDVTIELV 69
Qy 70 AHQGESALFNEVAPGYFSRDGQVQGVQVLAQVTKGARDSHAGLYMMHLVGHQRNRQV 129
Db 70 TSEKTSIIFNHTPPGNSKDSQNLHIQVQVLAQVTKGARDSHAGLYMMHLVGHQRNRQV 129
Qy 130 TLEVSGAEPQSDPTGFWFVPAVTVAVFILLVALVWFAYRCRCQQ 176
Db 130 NLTVNAADRKQTEDLPVTKVPDKPTAVRTEVIIIAITATIIITIGVFWY----- 182
Qy 177 RREKKFFLEPQMKVA 192
Db 183 ----KQFPVAPQIQMS 194

RESULT 5

US-10-052-648A-35
; Sequence 35, Application US/10052648A
; Publication No. US2004000558A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David
; APPLICANT: Burgess, Catherine
; APPLICANT: Casman, Stacie
; APPLICANT: Colman, Steven
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Ellerman, Karen
; APPLICANT: Gerlach, Valerie
; APPLICANT: Gunther, Erik
; APPLICANT: Kekuda, Ramesh

; APPLICANT: MacDougall, John R.
; APPLICANT: Mehraban, Fuad
; APPLICANT: Patturajan, Meera
; APPLICANT: Rothenberg, Mark
; APPLICANT: Shimkets, Richard
; APPLICANT: Smithson, Glennda
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Stone, David J.
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Zerhusen, Bryan D.
; TITLE OF INVENTION: PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS OF
; FILE REFERENCE: 21402-250 (CURA-550)
; CURRENT APPLICATION NUMBER: US/10/052,648A
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: 60/262,454
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: 60/272,920
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: 60/284,549
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/303,229
; PRIOR FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 60/262,892
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/263,605
; PRIOR FILING DATE: 2001-01-23
; PRIOR APPLICATION NUMBER: 60/269,098
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: 60/264,159
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/271,855
; PRIOR FILING DATE: 2001-02-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 35
; LENGTH: 969
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (848)..(889)
; OTHER INFORMATION: Where Xaa is any amino acid
US-10-052-648A-35
Query Match 7.5%; Score 98; DB 15; Length 969;
Best Local Similarity 24.9%; Pred. No. 1.4;
Matches 51; Conservative 24; Mismatches 86; Indels 44; Gaps 10;
Qy 2 QTCPLAFPG---HVSQALGTLFLAASLSAQN-----EGWDSPTCTGVSWSGENT 51
Db 620 QACPPGFWGPAFCACSHN-----GASCSAEDGACHCTPGWTGLFQRCPAFFGKDC 674
Qy 52 VMSCNISNAFS--HVNILKRAH-----QGESALFNEVAPGYFSRDGQVQVLAQVTKGARDSHAGLYMMHLVGHQRNRQVTLVSGAEPQSDPTGFWFVPAVTV 95
Db 675 GRVCCQCGASCDHISGKCTCTGFTGHCHE--QRCAPGTFGYCGCQLCECMNNSTCDHV 732
Qy 96 QGG-VAQVIVKIGARDSHAGLYMMHLVGHQRNRQVTLVSGAEPQSDPTGFWFVPAVTV 154
Db 733 TGTCCYCPGFGKIRCDQAALMMEELNPNYTKISPAL-----GAERHSV-----GAVTGMIL 782
Qy 155 AVFILLVALVWFAYRCRCQQREE 179
Db 783 LLFLIVVLLGLFAWHRRRRQKEKGD 807
RESULT 6
US-10-052-648A-34
; Sequence 34, Application US/10052648A
; Publication No. US2004000558A1

GENERAL INFORMATION:
APPLICANT: Anderson, David
APPLICANT: Burgess, Catherine
APPLICANT: Casman, Stacie
APPLICANT: Colman, Steven
APPLICANT: Edinger, Shlomit R.
APPLICANT: Ellerman, Karen
APPLICANT: Gerlach, Valerie
APPLICANT: Gunther, Erik
APPLICANT: Kekuda, Ramesh
APPLICANT: MacDougall, John R.
APPLICANT: Mehraban, Fuad
APPLICANT: Patterajan, Meera
APPLICANT: Rothenberg, Mark
APPLICANT: Shinkets, Richard
APPLICANT: Smithson, Glenda
APPLICANT: Spytek, Kimberly A.
APPLICANT: Stone, David J.
APPLICANT: Vernet, Corine A.M.
APPLICANT: Zerhusen, Bryan D.
TITLE OF INVENTION: PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS OF
FILE REFERENCE: 21402-250 (CURA-550)
CURRENT APPLICATION NUMBER: US/10/052,648A
CURRENT FILING DATE: 2002-12-09
PRIOR APPLICATION NUMBER: 60/262,454
PRIOR FILING DATE: 2001-01-18
PRIOR APPLICATION NUMBER: 60/272,920
PRIOR FILING DATE: 2001-03-02
PRIOR APPLICATION NUMBER: 60/284,549
PRIOR FILING DATE: 2001-04-18
PRIOR APPLICATION NUMBER: 60/303,229
PRIOR FILING DATE: 2001-07-05
PRIOR APPLICATION NUMBER: 60/262,892
PRIOR FILING DATE: 2001-01-19
PRIOR APPLICATION NUMBER: 60/263,605
PRIOR FILING DATE: 2001-01-23
PRIOR APPLICATION NUMBER: 60/269,098
PRIOR FILING DATE: 2001-02-15
PRIOR APPLICATION NUMBER: 60/264,159
PRIOR FILING DATE: 2001-01-25
PRIOR APPLICATION NUMBER: 60/265,517
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/271,855
PRIOR FILING DATE: 2001-02-27
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 97
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 34
LENGTH: 969
TYPE: PRT
ORGANISM: Homo sapiens
US-10-052-648A-34

Query Match 7.3%; Score 96; DB 15; Length 969;
Best Local Similarity 24.9%; Pred. No. 2.2;
Matches 51; Conservative 23; Mismatches 87; Indels 44; Gaps 10;

QY 2 QTCPLAFPG-----HVSQALGTLFLAASLSAQN-----EGWDSPICTEGVSVSGENT 51
DB 620 QACPPGFNGPACFACSCN-----GASCSAEDGACHCTPGWGLFCTQRCPPAFAFGKDC 674

QY 52 VMSCNISNAFS--HVNKILRAH-----GQSAIFNEVAPGYSRDGWQL-----QV 95
DB 675 GRVCOQNGASCDHLSGKCTCTGTGTGQHC--QRCAPGTGYGQQQLCECMNNSTCDHV 732

QY 96 QGG-VAQLVIKGARDSHAGLYNMHLVGHORNNRQVTLVEGSAEPQSPADTCGFWDVPAVVT 154
DB 733 TGTCTCSPGFGIRCDQQAALNMEELNPTYTKISPAL-----GAERHSV-----GAVTGIML 782

QY 155 AVFILLVALVMEAWYRCRCQORRE 179
DB 783 LLFFIVLLGLFAWHRRRQKGRD 807

RESULT 7
US-10-072-012-446
Sequence 446, Application US/10072012
Publication No. US20040033493A1
GENERAL INFORMATION:
APPLICANT: Tchernev, Velizar
APPLICANT: Spytek, Kimberly
APPLICANT: Zerhusen, Bryan
APPLICANT: Patturajan, Meera
APPLICANT: Shinkets, Richard
APPLICANT: Li, Li
APPLICANT: Gangolli, Esha
APPLICANT: Padigaru, Muralidhara
APPLICANT: Anderson, David W.
APPLICANT: Rastelli, Luca
APPLICANT: Miller, Charles E.
APPLICANT: Gerlach, Valerie
APPLICANT: Taupier Jr, Raymond J.
APPLICANT: Gusev, Vladimir Y.
APPLICANT: Colman, Steven D.
APPLICANT: Wolenc, Adam R.
APPLICANT: Pena, Carol E. A
APPLICANT: Furtak, Katarzyna
APPLICANT: Grosse, William M.
APPLICANT: Alsobrook II, John P.
APPLICANT: Lepley, Denise M.
APPLICANT: Rieger, Daniel K.
APPLICANT: Burgess, Catherine E.
TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-258
CURRENT APPLICATION NUMBER: US/10/072,012
CURRENT FILING DATE: 2002-01-31
PRIOR APPLICATION NUMBER: 60/265,102
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 60/265,514
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,517
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,412
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,395
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/266,406
PRIOR FILING DATE: 2001-02-02
PRIOR APPLICATION NUMBER: 60/266,767
PRIOR FILING DATE: 2001-02-05
PRIOR APPLICATION NUMBER: 60/267,057
PRIOR FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/266,975
PRIOR FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/267,459
PRIOR FILING DATE: 2001-02-08
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1391
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 446
LENGTH: 506
TYPE: PRT
ORGANISM: Bos taurus
US-10-072-012-446

Query Match 7.3%; Score 95.5; DB 15; Length 506;
Best Local Similarity 26.6%; Pred. No. 1;
Matches 51; Conservative 25; Mismatches 79; Indels 37; Gaps 8;

QY 4 CPLAFPGHVSQALGTLFLAASLSAQN-----EGWDSPICTEGVSVSGENTVMSCNISNAF-S 62
DB 228 CEVA---HVTLQGGPPLRGNTANLS---ETIRVPPTLEITRSPSAGNQVNTCOVNFYPR 281

QY 63 HVNKLRAHG-----QSAIFNEVAPGYSRDGWQLQVGGVAQLVIKGARDSHAGLYNM 117

Db 282 HLQTLWLNENGRTEAASVLVENKQDFTNQTSMLLVNSSAHREAVLTCQ----- 332
Qy 118 HLHGQRN---NRQVTLVSGAEPOSADPTGFWPVP-----AVVTAVFILLVALVM 165
Db 333 --VEHDROPVSKNHTLEVSA--PQKQDQGTGTPCPNDNNWTSIFIVGVVCCALLVALLI 388
Qy 166 PAWTRCSCQOR 177
Db 389 AALYLLRIRQNK 400
RESULT 8
US-10-425-114-62807
; Sequence 62807, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425.114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 62807
; LENGTH: 597
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3279-308-G7_FLI.pep
US-10-425-114-62807
Query Match 7.0%; Score 91.5; DB 15; Length 597;
Best Local Similarity 19.9%; Pred. No. 3.3;
Matches 62; Conservative 44; Mismatches 84; Indels 121; Gaps 16;
Qy 21 FLAASLSAQNGWSP-----ICTEGVSVSW-----GE 49
Db 299 FKAASAAAGHPDWELPDDAGINDPTDPTGFTTAERGTYLTEQGRFFLTWYSRKLIOHGD 358
Qy 50 NTVMSCSNISAFSHVNIKLRA-----HGOESALFNEVAPGYFS---RDGWLQVQ 96
Db 359 RVLEAN--KAPLGCKVKLAAKVSGIHWYRHPSHAA---ELTSGYVNLGGRDGY----- 408
Qy 97 GGVAQLVTKGARDSHAGLYMHLVGHQRNRRQVTLVSGAEPQAPDPTGFWPVPVAVTAV 156
Db 409 APIARMLAR-----HDGAVLNTCAEMNSQ-----AEEALSAPEQ----- 445
Qy 157 FILLVALVYFAWYR-----CRCQOREKFF-----LLEPQMKVAALBAGAQOGLS----- 203
Db 446 ---LVQVQLAGWREGVEVACENALSYDRRGYNQMLNARPNVGLSGAARRVAATF 502
Qy 204 -RASAEL-----WTPDSETPRPLALVFKPSPGLGALELL-----S 237
Db 503 LRLSDELLASNNFRIFRFRVVKMHADLDYCPDADRYGRPLKPLERSAPEMPMLLEATA 562
Qy 238 POPLFFPYAADP 248
Db 563 PAPAPFP--DP 571

RESULT 9
US-10-080-334-290
; Sequence 290, Application US/10080334
; Publication No. US20040002584A1
; GENERAL INFORMATION:
; APPLICANT: Pena, Carol E. A.
; APPLICANT: Shimkets, Richard A
; APPLICANT: Li, Li

; APPLICANT: Shenoy, Sureah G
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Spvtek, Kimberly A
; APPLICANT: Vernet, Corine A. M.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Guo, Xiaojia
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Casman, Stacie J
; APPLICANT: Boldog, Ferenc L
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Patturajan, Meera
; APPLICANT: Gangolli, Esha A
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Liu, Xiaohong
; APPLICANT: Baumgartner, Jason C.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Spaderna, Steven K
; APPLICANT: Zerhusen, Bryan D
; TITLE OF INVENTION: Proteins, Polynucleotides Encoding Them and Methods of
; FILE REFERENCE: 21402-275
; CURRENT APPLICATION NUMBER: US/10/080.334
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: 60/270,523
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: 60/322,712
; PRIOR FILING DATE: 2001-09-17
; PRIOR APPLICATION NUMBER: 60/311,980
; PRIOR FILING DATE: 2001-08-13
; PRIOR APPLICATION NUMBER: 60/330,307
; PRIOR FILING DATE: 2001-10-18
; PRIOR APPLICATION NUMBER: 60/278,796
; PRIOR FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: 60/281,521
; PRIOR FILING DATE: 2001-04-04
; PRIOR APPLICATION NUMBER: 60/276,677
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/311,595
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/270,220
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: 60/274,295
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/318,526
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/286,548
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: 60/291,765
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 60/270,797
; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 60/276,400
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/270,810
; PRIOR FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 388
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 290
; LENGTH: 6620
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-080-334-290
Query Match 7.0%; Score 91; DB 15; Length 6620;
Best Local Similarity 22.8%; Pred. No. 92;
Matches 58; Conservative 34; Mismatches 102; Indels 60; Gaps 10;
Qy 11 HVSQLGTLLFLAASLSAQNGWDSPTCTEGV-----VSVWCGENTVMSCNISNA 60
Db 1779 HGAQVLDLSAIY---SCRVGAEQDFPQVVEVAAKFCLELFPVCGELGGTTLACELSPA 1835
Qy 61 FSHVNIKLRAHQBSAIFNEVAPGYFSRDGWLQVQGVQVAFYKRGARDSHAGLYMHLV 120

TITLE OF INVENTION: 59079 and 12599, Protein Kinase Family
TITLE OF INVENTION: Members and Uses Therefor
FILE REFERENCE: MPI2001-047PIRCF1(M)
CURRENT APPLICATION NUMBER: US/10/077,130
PRIOR FILING DATE: 2002-02-15
PRIOR APPLICATION NUMBER: 60/269201
PRIOR FILING DATE: 2001-02-15
NUMBER OF SEQ ID NOS: 9
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5
LENGTH: 7968
TYPE: PRT
ORGANISM: Homo sapiens
US-10-077-130-5

Db 1836 CAEVVWRC-----GNTQPRVGR--FQVVAEGPVRSLTVGLRAEDAGEY----V 1879
Qy 121 GHORNNR---QVTLV-----SCAEPQAPDTGFWPVPVAVTAVFILLVALVMPAWYR 170
Db 1880 CESRDDHTSAQITVSVPRVWKFMSGLSTVVAEGGEATFCQVSPSDVAVV-----WFR 1933
Qy 171 CRCSQORREKFFLLEPQMKVAALRAGAQQGLSRASAEI-----WTPDSEPTPRPLALVF 225
Db 1934 DGA-----LLQPSKFAISQSGASHSLTISDLVLEDAGQITVEAGASSAALRV 1983
Qy 226 KPSPLGALELLSPQ 239
Db 1984 REAPVLFKKLEPQ 1997

Query Match 7.0%; Score 91; DB 13; Length 7968;
Best Local Similarity 22.8%; Pred. No. 1.2e+02;
Matches 58; Conservative 34; Mismatches 102; Indels 60; Gaps 10;

Qy 11 HVSQALGTLLFLAASLSAQNEGWDSPTCTEGV-----VSVSGENTVMSCNISNA 60
Db 1779 HGAQVLDLSAIY---SCRVGAEGQDPFVQVEEVAAKFCRLLEPVCGLGTTVLACELSPA 1835
Qy 61 FSHVNIKLRAHQESAI FNEVAPGYFRDGMQLQVGGVAQLVIKGARDSHAGLYMHLV 120
Db 1836 CAEVVWRC-----GNTQPRVGR--FQVVAEGPVRSLTVGLRAEDAGEY----V 1879
Qy 121 GHORNNR---QVTLV-----SCAEPQAPDTGFWPVPVAVTAVFILLVALVMPAWYR 170
Db 1880 CESRDDHTSAQITVSVPRVWKFMSGLSTVVAEGGEATFCQVSPSDVAVV-----WFR 1933
Qy 171 CRCSQORREKFFLLEPQMKVAALRAGAQQGLSRASAEI-----WTPDSEPTPRPLALVF 225
Db 1934 DGA-----LLQPSKFAISQSGASHSLTISDLVLEDAGQITVEAGASSAALRV 1983
Qy 226 KPSPLGALELLSPQ 239
Db 1984 REAPVLFKKLEPQ 1997

RESULT 12
US-09-759-130B-275
Sequence 275, Application US/09759130B
Publication No. US20030022279A1
GENERAL INFORMATION:
APPLICANT: Millennium Pharmaceuticals, Inc.
APPLICANT: McCarthy, Sean A
APPLICANT: Fraser, Christopher C
APPLICANT: Sharp, John D
APPLICANT: Barnes, Thomas S
APPLICANT: Kirt, Susan J
APPLICANT: Mackay, Charles R
APPLICANT: Myers, Paul S
APPLICANT: Leiby, Kevin R
APPLICANT: Wrighton, Nicolas
APPLICANT: Goodearl, Andrew
APPLICANT: Holtzman, Douglas A
TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
FILE REFERENCE: MPI00-5350NNIM
CURRENT APPLICATION NUMBER: US/09/759,130B
CURRENT FILING DATE: 2002-09-16
PRIOR APPLICATION NUMBER: US 09/479,249
PRIOR FILING DATE: 2000-01-07
PRIOR APPLICATION NUMBER: US 09/559,497
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: US 09/578,063
PRIOR FILING DATE: 2000-05-24
PRIOR APPLICATION NUMBER: US 09/333,159
PRIOR FILING DATE: 1999-06-14
PRIOR APPLICATION NUMBER: US 09/596,194
PRIOR FILING DATE: 2000-07-14

Query Match 7.0%; Score 91; DB 16; Length 6620;
Best Local Similarity 22.8%; Pred. No. 92;
Matches 58; Conservative 34; Mismatches 102; Indels 60; Gaps 10;
Qy 11 HVSQALGTLLFLAASLSAQNEGWDSPTCTEGV-----VSVSGENTVMSCNISNA 60
Db 1779 HGAQVLDLSAIY---SCRVGAEGQDPFVQVEEVAAKFCRLLEPVCGLGTTVLACELSPA 1835
Qy 61 FSHVNIKLRAHQESAI FNEVAPGYFRDGMQLQVGGVAQLVIKGARDSHAGLYMHLV 120
Db 1836 CAEVVWRC-----GNTQPRVGR--FQVVAEGPVRSLTVGLRAEDAGEY----V 1879
Qy 121 GHORNNR---QVTLV-----SCAEPQAPDTGFWPVPVAVTAVFILLVALVMPAWYR 170
Db 1880 CESRDDHTSAQITVSVPRVWKFMSGLSTVVAEGGEATFCQVSPSDVAVV-----WFR 1933
Qy 171 CRCSQORREKFFLLEPQMKVAALRAGAQQGLSRASAEI-----WTPDSEPTPRPLALVF 225
Db 1934 DGA-----LLQPSKFAISQSGASHSLTISDLVLEDAGQITVEAGASSAALRV 1983
Qy 226 KPSPLGALELLSPQ 239
Db 1984 REAPVLFKKLEPQ 1997

RESULT 11
US-10-077-130-5
Sequence 5, Application US/10077130
Publication No. US20020168742A1
GENERAL INFORMATION:
APPLICANT: Kapeller-Libermann, Rosana
APPLICANT: Acton, Susan L.

;; PRIOR APPLICATION NUMBER: US 09/342,364
;; PRIOR FILING DATE: 1999-06-29
;; PRIOR APPLICATION NUMBER: US 09/608,452
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/393,996
;; PRIOR FILING DATE: 1999-09-10
;; PRIOR APPLICATION NUMBER: US 09/602,871
;; PRIOR FILING DATE: 2000-06-23
;; PRIOR APPLICATION NUMBER: US 09/420,707
;; PRIOR FILING DATE: 1999-10-19
;; NUMBER OF SEQ ID NOS: 460
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 275
;; LENGTH: 435
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-759-130B-275

Query Match 6.8%; Score 89; DB 10; Length 435;
Best Local Similarity 31.0%; Pred. No. 3.8;
Matches 44; Conservative 18; Mismatches 62; Indels 18; Gaps 9;
QY 82 APGYFSDGWLQVGGVAQLVKGARDSHAGLYMHLVG---HQRNRRQVTLVSGAEP 138
DB 301 APG--SRDGSIAVLADG--SLAIGNVQEQHAGLFVCLATGPRLHHNQTHEYNVSVHFFRP 356
QY 139 Q-SAPDTGFWPVPVAVTAVFILLVLMFAMVYRCRCQORREKFFLLEPQMKVAALRAG 197
DB 357 EPEAFNTGF--TTLLGCAVGLVLLYLFA--PPCRCC--RRACPLPLAPNTQAP-RAE 410
QY 198 AQOGLSRASAEIWTDPSEPTPR 219
DB 411 PH----KSSVLSTTPPDAPSPQ 428

RESULT 13
US-10-189-123-5
;; Sequence 5, Application US/10189123
;; Publication No. US20030082586A1
;; GENERAL INFORMATION:
;; APPLICANT: KIRST, Susan J.
;; APPLICANT: HOLTZMAN, Douglas A.
;; APPLICANT: FRASER, Christopher C.
;; APPLICANT: SHARP, John D.
;; APPLICANT: BARNES, Thomas S.
;; TITLE OF INVENTION: ANTIBODIES HAVING DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
;; FILE REFERENCE: 10147-11U3
;; CURRENT APPLICATION NUMBER: US/10/189,123
;; CURRENT FILING DATE: 2002-07-02
;; PRIOR APPLICATION NUMBER: US 09/596,194
;; PRIOR FILING DATE: 2000-06-16
;; PRIOR APPLICATION NUMBER: US 09/342,364
;; PRIOR FILING DATE: 1999-06-29
;; NUMBER OF SEQ ID NOS: 100
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 5
;; LENGTH: 435
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-189-123-5

Query Match 6.8%; Score 89; DB 14; Length 435;
Best Local Similarity 31.0%; Pred. No. 3.8;
Matches 44; Conservative 18; Mismatches 62; Indels 18; Gaps 9;
QY 82 APGYFSDGWLQVGGVAQLVKGARDSHAGLYMHLVG---HQRNRRQVTLVSGAEP 138
DB 301 APG--SRDGSIAVLADG--SLAIGNVQEQHAGLFVCLATGPRLHHNQTHEYNVSVHFFRP 356
QY 139 Q-SAPDTGFWPVPVAVTAVFILLVLMFAMVYRCRCQORREKFFLLEPQMKVAALRAG 197
DB 357 EPEAFNTGF--TTLLGCAVGLVLLYLFA--PPCRCC--RRACPLPLAPNTQAP-RAE 410

QY 198 AQOGLSRASAEIWTDPSEPTPR 219
DB 411 PH----KSSVLSTTPPDAPSPQ 428
RESULT 14
US-10-188-495-5
;; Sequence 5, Application US/10188495
;; Publication No. US20030175733A1
;; GENERAL INFORMATION:
;; APPLICANT: KIRST, Susan J.
;; APPLICANT: HOLTZMAN, Douglas A.
;; APPLICANT: FRASER, Christopher C.
;; APPLICANT: SHARP, John D.
;; APPLICANT: BARNES, Thomas S.
;; TITLE OF INVENTION: POLYPEPTIDES HAVING DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
;; FILE REFERENCE: 10147-11U2
;; CURRENT APPLICATION NUMBER: US/10/188,495
;; CURRENT FILING DATE: 2002-07-02
;; PRIOR APPLICATION NUMBER: US 09/596,194
;; PRIOR FILING DATE: 2000-06-16
;; PRIOR APPLICATION NUMBER: US 09/342,364
;; PRIOR FILING DATE: 1999-06-29
;; NUMBER OF SEQ ID NOS: 100
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 5
;; LENGTH: 435
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-188-495-5

Query Match 6.8%; Score 89; DB 14; Length 435;
Best Local Similarity 31.0%; Pred. No. 3.8;
Matches 44; Conservative 18; Mismatches 62; Indels 18; Gaps 9;
QY 82 APGYFSDGWLQVGGVAQLVKGARDSHAGLYMHLVG---HQRNRRQVTLVSGAEP 138
DB 301 APG--SRDGSIAVLADG--SLAIGNVQEQHAGLFVCLATGPRLHHNQTHEYNVSVHFFRP 356
QY 139 Q-SAPDTGFWPVPVAVTAVFILLVLMFAMVYRCRCQORREKFFLLEPQMKVAALRAG 197
DB 357 EPEAFNTGF--TTLLGCAVGLVLLYLFA--PPCRCC--RRACPLPLAPNTQAP-RAE 410
QY 198 AQOGLSRASAEIWTDPSEPTPR 219
DB 411 PH----KSSVLSTTPPDAPSPQ 428

RESULT 15
US-10-741-790-275
;; Sequence 275, Application US/10741790
;; Publication No. US20040121396A1
;; GENERAL INFORMATION:
;; APPLICANT: Millennium Pharmaceuticals, Inc.
;; APPLICANT: McCarthy, Sean A
;; APPLICANT: Fraser, Christopher C
;; APPLICANT: Sharp, John D
;; APPLICANT: Barnes, Thomas S
;; APPLICANT: Kirst, Susan J
;; APPLICANT: Mackay, Charles R
;; APPLICANT: Myers, Paul S
;; APPLICANT: Leiby, Kevin R
;; APPLICANT: Wrighton, Nicolas
;; APPLICANT: Goodearl, Andrew
;; APPLICANT: Holtzman, Douglas A
;; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
;; TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
;; TITLE OF INVENTION: USES
;; FILE REFERENCE: MPI00-5350MIM
;; CURRENT APPLICATION NUMBER: US/10/741,790
;; CURRENT FILING DATE: 2003-12-19
;; PRIOR APPLICATION NUMBER: US 09/479,249
;; PRIOR FILING DATE: 2000-01-07

```

; PRIOR APPLICATION NUMBER: US 09/559,497
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 09/578,063
; PRIOR FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: US 09/333,159
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: US 09/596,194
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/342,364
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: US 09/608,452
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/393,996
; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: US 09/602,871
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 09/420,707
; PRIOR FILING DATE: 1999-10-19
; NUMBER OF SEQ ID NOS: 460
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 275
; LENGTH: 435
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-741-790-275

Query Match      6.8%; Score 89; DB 16; Length 435;
Best Local Similarity 31.0%; Pred. No. 3.8; Indels 18; Gaps 9;
Matches 44; Conservative 18; Mismatches 62;

QY      82  AFGYERDGMQLQVGGVAQVLTKGARDSHAGLYMHLVG---HORNRRQVTLVSGAEP 138
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      301  AFG--SRDGSIAVLADG--SLAIGNVOEQHAGLVCLATGPRLLHNOHEYNVSVHPPRP 356
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY      139  Q-SAPDQFWPVPVAVTAVFILLVALVMFAYWRCRCQORREKKFLLPEQMKVAALEAG 197
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      357  EEEAENTGCF--TTLLGCAVGLVLLYLFA--PPCRCC--RRACPLPPLAPNTQAP--RAE 410
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY      198  AQQGLSRASAEIWTDPDSEPTPR 219
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      411  PH-----KSVLSLTPTPDAPSPQ 428
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Search completed: March 20, 2005, 10:15:10
Job time : 77 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 20, 2005, 09:17:28 ; Search time 25 Seconds
(without alignments)
954.470 Million cell updates/sec

Title: US-10-080-522-1
Perfect score: 1308
Sequence: 1 MQTCPLAFPGHVSQLGTL.....PLGALELLSQPLFFPYAADP 248
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 9621673 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	103.5	7.9	550	S64725	probable lipoprote
2	94.5	7.2	393	T07167	probable isocitrat
3	90.5	6.9	271	T36985	probable membrane
4	90	6.9	1088	H96747	unknown protein fl
5	89	6.8	798	A12053	competence protein
6	86.5	6.6	929	A44048	genome polypotein
7	86	6.6	493	S51574	mocR protein - Rhi
8	85.5	6.5	420	A82699	conserved hypothet
9	85.5	6.5	468	B46114	glycoprotein gp13
10	85.5	6.5	2303	GNNTTP	genome polypotein
11	85.5	6.5	2303	S13554	genome polypotein
12	84.5	6.5	289	A83608	polyamine transpor
13	83.5	6.4	468	1 VGBE18	glycoprotein gp13
14	83	6.3	550	1 VGBE18	glycoprotein E - h
15	83	6.3	436	E71086	hypothetical prote
16	82.5	6.3	233	T15620	hypothetical prote
17	82.5	6.3	288	A55737	PD-1 protein - hum
18	82.5	6.3	636	F72736	hypothetical prote
19	82	6.3	398	T46475	hypothetical prote
20	81	6.2	450	1 FOLJFP	gag polypotein -
21	81	6.2	450	S23819	gag protein - feli
22	80.5	6.2	312	C87562	conserved hypothet
23	80.5	6.2	541	T48811	hypothetical prote
24	80.5	6.2	1240	T03097	CDO protein - huma
25	80.5	6.2	2301	GNNTYN	genome polypotein
26	80.5	6.2	2303	1 GNNTYM	genome polypotein
27	80	6.1	290	T09260	aquaporin-like tra
28	79	6.0	407	B70962	hypothetical prote
29	79	6.0	950	T51134	ionotropic glutama

ALIGNMENTS

RESULT 1

S64725
probable lipoprotein uxpA precursor - Pseudomonas putida
C:Species: Pseudomonas putida
C>Date: 06-Dec-1996 #sequence_revision 06-Dec-1996 #text_change 09-Jul-2004
C:Accession: S64725; S47503
R:de Groot, A.; Krijger, J.J.; Filloux, A.; Tommassen, J.
Mol. Gen. Genet. 250, 491-504, 1996
A:Title: Characterization of type II protein secretion (xcp) genes in the plant growth-
A:Reference number: S64724; MUID:96186881; PMID:8602167
A:Accession: S64725
A:Molecule type: DNA
A:Residues: 1-550 <DEG>
A:Cross-references: UNIPROT:Q52289; EMBL:X81085; NID:g3293032; PIDN:CAA56977.1; PID:g553
C:Genetics:
A:Gene: uxpA
F:1-23/Domain: signal sequence #status predicted <SIG>
F:30-550/Product: probable lipoprotein uxpA #status predicted <MAT>

Query Match 7.9%; Score 103.5; DB 2; Length 550;
Best Local Similarity 23.2%; Pred. No. 0.17;
Matches 72; Conservative 42; Mismatches 118; Indels 79; Gaps 16;

QY	2	QTCPLAFPGHVSQLGTLFLAA-----SLSAQN-EGWDSPTCTEGVVSVS	46
DB	90	QLAFLDAGRADEANLGGYAVLGALLEQLRGAGAGNSLTLENGQWNG-----SGLAYLT	145
QY	47	WGENTVMGNCISNAFSHVNIKLRA--HGQESAIFNEVAPGYFSR-----DGWQ	92
DB	146	QGESGVQSGQLLGSEARVSSDERVLWPQPSAALYRQASATTLGAGLADQRQALGLEPLQ	205
QY	93	LQVQGVQALVTKG-----ARDSHAGLYNMWH---LVGHQNNRVQTVLEVSAGAPQAP--	142
DB	206	LFERGG-ARIAVGVGTVDPYAQDKASLKQMYQALLPVFOQARREADLVVALADVGTGPG	264
QY	143	-----DTGFVPVPAVTVTAFTLLVALVMFAW-----YRCRC-----SQ	176
DB	265	WLAEPLPAIDLLLCARGQDLWETPVQATQASGRVP-VLFACRGSGAFRLRCQVAGQW	323
QY	177	REEKXFF-----LLBFPQMKVAALRAGAQQGLSRASAEMLWTPDSEPTPR-PLALVF	230
DB	324	QPEGRFFPTFEQTLSPAAQVRAGLQALNQQRAGHAAML--DQPLARAPQALWRRDTRG	381
QY	231	GALELLSQPL 241	
DB	382	GSWDRLLHQAL 392	

RESULT 2

T07167
probable isocitrate dehydrogenase (NAD) (EC 1.1.1.41) chain 1 - tomato
C:Species: Lycopersicon esculentum (tomato)

Db 228 GVKGVTSPL-----PGPSGNVE 245

RESULT 4

H96747

unknown protein T10D10.14 [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004

C;Accession: H96747

R;Theologas, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C. C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A;Authors: Salzberg, S.B.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A;Reference number: A86141; MUID:21016719; PMID:11130712

A;Accession: H96747

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-1088 <STO>

A;Cross-references: UNIPROT:Q9C9D7; GB:AE005173; NID:g6730762; PIDN:AAF27151.1; GSPDB:G

C;Genetics:

A;Gene: T10D10.14

A;Map position: 1

Query Match 6.9%; Score 90; DB 2; Length 1088;

Best Local Similarity 21.2%; Pred. No. 6.3; Indels 66; Gaps 11;

Matches 55; Conservative 46; Mismatches 93;

QY 15 ALGTLFLAASLSAQNEGWDSPICTEGVSVSGENTVMSCTNSINAFSHVNIKLRAHG-- 72

DB 789 ALGNQLNAGR-GMGGTGISSMSVPGIGN--QGNP-MNLNPASNLNAISQQLRSALT 844

QY 73 -QESAFNEVAPGYFSDGWLQVGGVAQLVIKGDHSH---AGLYMHLVGHORNRQ 128

DB 845 PQNALFTQIRMGANRGVNGAPQTGISG--VSGTRQMPHSSAGLSM----- 890

QY 129 VTLEVSGAEPOSADPTGFWPVPVAVTAVFILLVALVMPFAYRCRCQORREKFFLLEPQ 188

DB 891 --LDQNRANLQRAAMGNMGPPKLMPGMNLVM-----NQOQQOQLQQOPQ 935

QY 189 MKVAALRAGAQQLSRASAEI-WTP-----DSEPTPRPLALYFKPSPL 230

DB 936 QQQLQHQQQLQPMSPQSQQLAQSPQQOQLQHQHQPQQOQQOATASPLQSVLSPFPQ 995

QY 231 GA-----LELLSPQL 241

DB 996 GSPSAGITQQQLQQSSPQOM 1015

RESULT 5

AI2053

competence protein [imported] - Nostoc sp. (strain PCC 7120)

C;Species: Nostoc sp. PCC 7120

A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004

C;Accession: AI2053

R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, DNA Res. 8, 2005-213, 2001

A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana

A;Reference number: AB1807; MUID:21595285; PMID:11759840

A;Accession: AI2053

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-798 <KUR>

A;Cross-references: UNIPROT:Q9YV14; GB:BA000019; PIDN:BA073682.1; PID:g17131073; GSPDB:G

A;Experimental source: strain PCC 7120

C;Genetics:

Query Match 6.9%; Score 90.5; DB 2; Length 271;

Best Local Similarity 25.1%; Pred. No. 1.1; Indels 8;

Matches 51; Conservative 21; Mismatches 60;

QY 83 PGYFSRDGWLQVGGVAQ-----LVIKGAR-----DSH 111

DB 63 PIYVSGGHKL--AGALAAFPVHGLVVEGRALDAGASTGGFTDVLRLAGAAHVAVVDVG 120

QY 112 AGLYMHL-----VGHORNRQVTLVSGAEPOSADPTGFWPVPVAVTAVFILLVALV 164

DB 121 YGQLAWSLRQDRVTVKORTNRELTPEAIDG-----PVDLVGDLSTPIPLAV 170

QY 165 MFAWYRCRCQORREKFFLLEPQMKV-----RALRAGAQQLSRASAEIWT 211

DB 171 LPALVRC---TRPGADLVNMFQPEVGVKRLGSGGVRSQALRAEAVRGVARKAWELGL 227

QY 212 PDSEPTPRPLALYFKPSPLGAL 234

C;Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 09-Jul-2004

C;Accession: T07167

R;Tahiri-Alaoui, A.; Avrova, A.; Antoniw, J.F.

submitted to the EMBL Data Library, January 1998

A;Description: A putative mitochondrial isocitrate dehydrogenase from tomato roots up-te

A;Reference number: Z15974

A;Accession: T07167

A;Status: translated from GB/EMBL/DBDJ

A;Molecule type: mRNA

A;Residues: 1-393 <ANT>

A;Cross-references: UNIPROT:082004; EMBL:Y16126; PIDN:CAA76076.1

A;Experimental source: cultivar Early Mech; root

C;Superfamily: 3-isopropylmalate dehydrogenase

C;Keywords: oxidoreductase

Query Match 7.2%; Score 94.5; DB 2; Length 393;

Best Local Similarity 23.8%; Pred. No. 0.76; Indels 65; Gaps 9;

Matches 51; Conservative 30; Mismatches 68;

QY 63 HVNKLRAHQESAIFNEVAPG---YFSRDGWLQVGGVAQLVIKGAR--DSHAGL-Y 115

DB 94 HGMKEYASGDVIGISN---PGRNKVFKREGKTPVGGVSSLVNQLRKELDLVSLVH 150

QY 116 MHLVGHQRNRQVTL-----EVSGAEPOSAP-----DT 144

DB 151 CFNLKGLPTRHENVDIWIWIRENTEGEYSGLHEHVEVGVESLKVMTKFCSERIAKYAF 210

QY 145 GFWPVPVAVTAVF---ILLVALVMPFAYRCRCQORREKFFLLEPQMKVVAALRAGAQG 201

DB 211 AYLNRRKVTVAVKANIMENLMLVYFESC-----RDRKQITLG----- 250

QY 202 LSRASAEIWTDPSEPTPRPLALYFKPSPLGAL 235

DB 251 -SSTMRELTTVACNLYQSLSLNMLMSPLISTEI 283

RESULT 3

T36885

probable membrane protein - Streptomyces coelicolor

C;Species: Streptomyces coelicolor

C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004

C;Accession: T36885

R;Murphy, L.; Harris, D.; James, K.D.; Parkhill, J.; Bartell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, August 1999

A;Reference number: Z21617

A;Accession: T36885

A;Status: preliminary; translated from GB/EMBL/DBDJ

A;Molecule type: DNA

A;Residues: 1-271 <MUR>

A;Cross-references: UNIPROT:Q9S218; EMBL:AL109848; PIDN:CAB52846.1; GSPDB:GN00070; SCORE

A;Experimental source: strain A3(2)

C;Genetics:

A;Gene: SCORDB:SCI51.22C

C;Superfamily: hemolysin homolog yqxC

Query Match 6.9%; Score 90.5; DB 2; Length 271;

Best Local Similarity 25.1%; Pred. No. 1.1; Indels 8;

Matches 51; Conservative 21; Mismatches 60;

QY 83 PGYFSRDGWLQVGGVAQ-----LVIKGAR-----DSH 111

DB 63 PIYVSGGHKL--AGALAAFPVHGLVVEGRALDAGASTGGFTDVLRLAGAAHVAVVDVG 120

QY 112 AGLYMHL-----VGHORNRQVTLVSGAEPOSADPTGFWPVPVAVTAVFILLVALV 164

DB 121 YGQLAWSLRQDRVTVKORTNRELTPEAIDG-----PVDLVGDLSTPIPLAV 170

QY 165 MFAWYRCRCQORREKFFLLEPQMKV-----RALRAGAQQLSRASAEIWT 211

DB 171 LPALVRC---TRPGADLVNMFQPEVGVKRLGSGGVRSQALRAEAVRGVARKAWELGL 227

QY 212 PDSEPTPRPLALYFKPSPLGAL 234

A:Gene: comE

Query Match 6.8%; Score 89; DB 2; Length 798;
Best Local Similarity 21.6%; Pred. No. 5.4;
Matches 44; Conservative 33; Mismatches 73; Indels 54; Gaps 9;

QY 40 EGVSVSVCNTVMSCNINAFSHV-----NIKLRHGQSAIFNEVAPGYFGRDGMQ 92
DB 606 QGVNQIDWAIATDFORNNNDWLEVLQRLAIKFNFYATNKENSLADQAIPOI----- 658
QY 93 LQVQGVSAQLVTKGARDSHAGLYMHVGHQRNNR-----QVTVLEVSAGAEPOSAP---DT 144
DB 659 LQKQGIIVQLLPVG-QTINLGTVAQLINEQPMQLQMLQGLQSWLIVGVDEPKVEVERIMKA 717
QY 145 GFWPVPVAVTAVFILLVALVMEFAYRCRCQORREKKFFLEPQMKVA---ALRAGAQQG 201
DB 718 GCMPSQVIL-----W-----CNASLKDLYMMLKPOVALIASSGSLSTVLS 759
QY 202 LSRASAEI-----WTPDSE 215
DB 760 LSKTSTKVFVTAQDGAIQMPNGE 783

RESULT 6

A44048
genome polyprotein - Vilyuisk virus (strain V-1) (fragment)
C:Species: Vilyuisk virus
C:Date: 11-Feb-1993 #sequence_revision 11-Feb-1993 #text_change 07-Feb-1997
C:Accession: A44048
R:Pritchard, A.E.; Strom, T.; Lipton, H.L.
Virology 191, 469-472, 1992
A:Title: Nucleotide sequence identifies Vilyuisk virus as a divergent Theiler's virus.
A:Reference number: A44048; MUID:93033144; PMID:1413519
A:Accession: A44048
A:Molecule type: genomic RNA
A:Residues: 1-929 <PRI>
A:CROSS-references: GB:M94868
A:Superfamily: foot-and-mouth disease virus genome polyprotein
C:Keywords: polyprotein

Query Match 6.6%; Score 86.5; DB 2; Length 929;
Best Local Similarity 27.0%; Pred. No. 11;
Matches 27; Conservative 16; Mismatches 34; Indels 23; Gaps 5;

QY 61 FSHVNIKL--RAHGQSAIFN-EVAPGYFSRDGMQVQGGVAQLVKGARDSHAGLYM 117
DB 224 FSHIRVPLPHALAGEHGGVFGATLRHYLAKCGRVQVQCNASQF-----HAGSLIV 275
QY 118 HL-----VGHQRNNRQVTVLEVSAGAEPOSAPDTGF 146
DB 276 FLAPEFYTGTVATSGQEPNKVFLMDTTWQEPQAP-TGF 314

RESULT 7

S51574
mocr protein - Rhizobium meliloti
C:Species: Rhizobium meliloti
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: S51574; S43169
R:Rosbach, S.; Kuipa, D.A.; Rosbach, U.; de Bruijn, F.J.
Mol. Gen. Genet. 245, 11-24, 1994
A:Title: Molecular and genetic characterization of the rhizopine catabolism (mocABRC) gene
A:Reference number: S51569; MUID:95147842; PMID:7845353
A:Accession: S51574
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-493 <ROS>
A:CROSS-references: UNIPROT:P49309; EMBL:X78503; NID:G468758; PID:G468764
C:Genetics:
A:Gene: mocr
C:Superfamily: hypothetical protein b1439

Query Match 6.6%; Score 86; DB 1; Length 493;

Best Local Similarity 26.0%; Pred. No. 5.8;
Matches 53; Conservative 20; Mismatches 65; Indels 66; Gaps 11;

QY 60 AFSHVNIKLRAHGQSAIFNEVA--PGY-----FSRDGMQVQGGVAQLVKGARDSHA 112
DB 96 SLSSRGMAAQAQPRDRTIPDRIAHPGYPETKAPFFSTW-----AKLLKRRHARYSHE 147
QY 113 GLYMHVLY-GHQRNNRQVTVLEVSAGAEPOSAPDTGFWPVPAVTVAVFILLVALVMEFA---- 167
DB 148 DLYGYHVTGHPR-----LKAATAEYLRA-SRGVECAPEQIVVNGTQAALDILARMLV 200
QY 168 -----WYRCRCQORREKKFFLEPQMKVAALRAGAQQGLSRASAEI-----WTP 212
DB 201 DEGDCIW-----MEEPGY-----IGAQNLSLSAGAKLVPLPVERDGNLSL 239
QY 213 DSEPTPRPLALVFKPS-----PLGAL 233
DB 240 EDETRPSRLIFVTPSCQWPLGCL 263

RESULT 8

A82699
conserved hypothetical protein XF1304 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: A82699
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717; PMID:10910347
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: A82699
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-420 <SIM>
A:CROSS-references: UNIPROT:Q9PDS5; GB:AE003963; GB:AE003849; NID:G9106285; PIDN:AAF8411
A:Experimental source: strain 9a5c
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrier, J. as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S. submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohne-Hagemann, G.; Kemp, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigret, J.D.; Junqueira, M.L.; Madeira, H.M.P.; Marino, C.L.; Marques, M.V.; Martins, J. Chado, M.A.; Madeira, A.M.B.N.; Matekuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.; F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.; Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Sancelli, R.V.; Sawasak M.; Tsunako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF1304

Query Match 6.5%; Score 85.5; DB 2; Length 420;
Best Local Similarity 21.1%; Pred. No. 5.4;
Matches 59; Conservative 36; Mismatches 110; Indels 75; Gaps 11;

QY 10 GHVSQAL-GTLLFLAASLSAQNEGWDSPICTEGVSVSGWNTVMSCNINAFSHVNIKL 68
DB 33 GTTAQRLEASLVALSQGLDCEPMSNP--TGILSPDPAKAIGSSDITRVI----- 83
QY 69 RAHGQSAIFNEVAPGYFSRDGMQVQGGVAQLVKGARDSHAGLYMHLVGHQRNNRQ 128
DB 84 -----RLAPG--ENDLYKLSVADNVAAQAVASGRMSISQGHYALRLDRELDRRS 130
QY 129 VTVLEVSAGAEPOSAPDTGFWPVP--AVVTAVFI-LVALVMEFAYRCRCQORREKKFFLL 185
DB 131 KWMRVFGPLAAGGVAGLWKLPLWLDVATAATVGLLIGLL-----TQVTDHRAATR 180
QY 186 EPQMKVALRAG-----AQQGLSRASAEIWTDPDS 214
DB 181 EASEAALLAGFVATLVATLIGPINLNTVITASVVVLLPGMLTNAVNELSSQHWVSQT 240

Best Local Similarity 26.2%; Pred. No. 4.3;
Matches 60; Conservative 27; Mismatches 91; Indels 51; Gaps 12;

QY 1 MQTCPLAFPGHVSQAALGTILFLAASLSAQNEGWDSPICTEGVTVSVSGENTVMSCNISNA 60
DB :
100 MYTAPLVMPVEITLISLLLFVA---MAQLIGWPQ---ERGIWTI-WIAHT-----SFC 146

QY 61 FSHWIKLRANGQESAIENFVAGPYFSRDGQW---LQVGGVAAQLVIKGARDSHAGLYMW 117
DB :
147 SSVAVVVVSARLRRLDLSIEEAMDLGAKPKWVFLLITPIMAPISLAAGGMWSFA-LSLD 205

QY 118 HLYCHQRNNQVTILEVSG-----AEPSADPTGFVPVPAVTVATVIFILLVAL-VNMFAY 169
DB || :
206 DLV-----LASFVSGPGSTTLPMVEVFAVLKVKPEINAVASLIILSVLSLTFFAWY 257

QY 170 RCRCSQQRRKKFPLELPQMKAALRAGAQQGLSRASAELTWTPDSEPTP 218
DB :
258 FTQRAEER-----RRAIQAMMEETATD-WQKGS-PTP 288

RESULT 13
VBEEH
glycoprotein gp13 precursor - equine herpesvirus 1
N:Alternate names: glycoprotein gC
C:Species: equine herpesvirus 1
A>Note: host Equus caballus (domestic horse)
C>Date: 30-Jun-1989 #sequence revision 30-Jun-1989 #text_change 09-Jul-2004
C:Accession: A28149; A32980; H36796
R:Allen, G.P.; Coogle, L.D.
J. Virol. 62, 2850-2858, 1988
A>Title: Characterization of an equine herpesvirus type 1 gene encoding a glycoprotein (gC)
A:Reference number: A28149; UID:88275055; PMID:2455821
A:Accession: A28149
A:Molecule type: DNA
A:Residues: 1-468 <ALL>
A:Cross-references: UNIPROT:P12889; GB:M86664; NID:G330791; PIDN:AAB02451.1; PID:G330808
A:Experimental source: strain Kentucky T431
A>Note: The authors translated the codon ACA for residue 43 as Pro
R:Guo, P.; Goebel, S.; Davis, S.; Perkins, M.E.; Languet, B.; Desmettre, P.; Allen, G.; F.
J. Virol. 63, 4189-4198, 1989
A>Title: Expression in recombinant vaccinia virus of the equine herpesvirus 1 gene encoded by the equine herpesvirus 1 genome
A:Reference number: A32980; UID:89382761; PMID:2550665
A:Accession: A32980
A:Molecule type: DNA
A:Residues: 1-468 <GUO>
A:Cross-references: NID:G330791; PIDN:AAB02451.1; PID:G330808
R:Telford, E.A.R.; Watson, M.S.; McBride, K.; Davison, A.J.
submitted to Genbank, March 1992
A>Description: The DNA sequence of equine herpesvirus-1.
A:Reference number: A36805
A:Accession: H36796
A:Molecule type: DNA
A:Residues: 1-468 <TEL>
A:Cross-references: GB:M86664; NID:G330791; PIDN:AAB02451.1; PID:G330808
R:Telford, E.A.R.; Watson, M.S.; McBride, K.; Davison, A.J.
Virology 189, 304-316, 1992
A>Title: The DNA sequence of equine herpesvirus-1.
A:Reference number: A41931; UID:92295566; PMID:1318606
A:Contents: annotation; possible protein-coding frames
A>Note: neither amino acid nor nucleotide sequence is given
C:Genetics:
A:Gene: 16
C:Superfamily: herpesvirus glycoprotein F
C:Keywords: Glycoprotein; transmembrane protein
F:1-30/Domain: signal sequence #status predicted <SIG>
F:31-468/Product: glycoprotein gp13 #status predicted <MAT>
F:46,57,62,92,100,131,203,208,269/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 6.4%; Score 83.5; DB 1; Length 468;
Best Local Similarity 22.5%; Pred. No. 9.3;
Matches 42; Conservative 23; Mismatches 69; Indels 53; Gaps 7;

```

Qy 2 QTCPLAPPGHVSQALGTLLFLAASLSAQNEGWSP-----ICTEGVVSVSWGENTWMS 54
Db 34 QSTP-ATPTHTTNLTAAHAGAGSDNTNANGTESTHSHETTITCTKSLISVPYKSDVMN 92
Qy 55 CNTSNAPSHVNIKLRAHQGSAINEVAPG---VF-----SDRG 90
Db 93 CTTSSVGWVSYSEYRLIYLNQRTPESTGTPPDGEENYINHNATKDQDTLLFSTAERKKSRRG 152
Qy 91 WLQGV-----QGGVA-QLVIKGARDSHAGLYMMHLV---GHORNRQV 129
Db 153 GOLGVIPDRLPKROLFLNLPHLHTEGGTKFPLTIKISVDWRTAGIYVMSLYAKNGTLVNSTSV 212
Qy 130 TLEVSGA 136
Db 213 TVSTYNA 219

RESULT 14
VGBE18
glycoprotein E - human herpesvirus 1
N/Alternate names: US8
C/Species: human herpesvirus 1
C/Date: 17-Mar-1987 #sequence_revision 17-Mar-1987 #text_change 09-Jul-2004
C/Accession: A03733; A45696
R/McGeoch, D.J.; Dolan, A.; Donald, S.; Rixon, F.J.
J. Mol. Biol. 181, 1-13, 1985
A/Title: Sequence determination and genetic content of the short unique region
A/Reference number: A00656; MUID:85160822; PMID:2984429
A/Accession: A03733
A/Molecule type: DNA
A/Residues: 1-550 <MCG>
A/Cross-references: UNIPROT:P04488; GB:X02138; NID:G59865; PIDN:CAA26062.1; PID:
R/Georgopoulou, U.; Michaelidou, A.; Roizman, B.; Mavromara-Nazos, P.
J. Virol. 67, 3961-3968, 1993
A/Title: Identification of a new transcriptional unit that yields a gene product
A/Reference number: A45696; MUID:93287213; PMID:8389914
A/Accession: A45696
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 438-550 <GEO>
A/Cross-references: GB:S62895; NID:G386127; PIDN:AAB27080.1; PID:G386128
A/Experimental source: R35
A/Note: sequence extracted from NCBI backbone (NCBIN:133646, NCBI:P:133647)
C/Superfamily: herpesvirus glycoprotein E
C/Keywords: Glycoprotein
F:124,243,501/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 6.3%; Score 83; DB 1; Length 550;
Best Local Similarity 25.4%; Pred. No. 12;
Matches 53; Conservative 16; Mismatches 62; Indels 78; Gaps 11

Qy 80 EVAPGYFSRDGWLQVGGVAQLVIKGARDSHAGLYMM-----HLVGH-----QR 124
Db 330 EPVPG-----LAWQAASVNLFEFRDASPOHSGLYLCVVYNDHIHAWGHITITSTAAGY 381
Qy 125 NNRQV-----TLEVSGAESQSPDPTGFVPVAVTVFLLVALVMEFAYR 170
Db 382 RNAVVEOPLPQRGADLAEPHHPHVGVGPPHPTTHGALRLGAVMGAA-LLLSALGUSVWAC 440
Qy 171 CRCSQRRKKFFLLEPQMKAALRAGAQQGLSR-----ASAEI---WTPDSE----- 215
Db 441 MTCWRR-----AWRAVKSRASGKPTVIRVADSELYADWSSDSGERDQ 485

Qy 216 -----PFRPLALVFKPSPLGA-LELISP 238
Db 486 VPMIAPPERP-----DSPSTNGSGFEILSP 510

```

RESULT 15
E71086
hypothetical protein PH0954 - Pyrococcus horikoshii
C:Species: Pyrococcus horikoshii

Sun Mar 20 14:17:43 2005

us-10-080-522-1.rpr

C>Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 12-Jul-2004
C/Accession: E71086
R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekin
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi
DNA Res. 5, 55-76, 1998
A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic
A:Reference number: A71000; MUID:98344137; PMID:9679194
A/Accession: E71086
A/Status: Preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-4436 <KAW>
A/Cross-references: UNIPROT:O58659; GB:AP000004; NID:G3236131; PIDN:BA30051.1; PID:d103
A/Experimental source: strain OT3
A/Note: this accession replaces an interim accession for a sequence replaced by GenBank
C/Genetics:
A/Gene: PH0954

Query Match 6.3%; Score 83; DB 2; Length 4436;
Best Local Similarity 23.0%; Pred. No. 1.4e+02;
Matches 37; Conservative 24; Mismatches 58; Indels 42; Gaps 8;

QY 34 DSPICTEGVVS-----WGENTVMSCNISNAFSHVNIKLA-----HGQE 74
DB 300 DDPDLNNLAEGVWPGDIWENASINNLIPEFASINFKVRTTSKIPSAKVLLRNGVE 359

QY 75 SAI-----FNEVAPGYFSRDGWLQVGG--VAQLVIKGARDSHA-----GLYMHILV 120
DB 360 EKIEYLSFYNGIAEGEITS--W--LVQGGNYTLALLVEGKGIDINSNNIYLLGNYNFPLP 414

QY 121 GHQRNNRQVTL-----EVSGAEPOSAPDTGFWPVPVAVTAVF 157
DB 415 NFEVGNYSIDLPTCVDSTGEVRVNTSTANWSIPVRLTVY 455

Search completed: March 20, 2005, 09:49:42
Job time : 30 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 20, 2005, 06:56:18 ; Search time 92 Seconds
(without alignments)
1380.387 Million cell updates/sec

Title: US-10-080-522-1
Perfect score: 1308
Sequence: 1 MQTCLAPGHVSQALFTLL.....PLGALLELPQLFPFYAADP 248

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_03.*

1: uniprot_eprot.*

2: uniprot_treml.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1308	100.0	248	1 SCTM_HUMAN	Q8WV66 homo sapien
2	382	29.2	192	2 Q921W8	Q921W8 mus musculus
3	376.5	28.8	218	2 Q6P781	Q6P781 rattus norv
4	358	27.4	201	2 Q6AYS0	Q6AYS0 rattus norv
5	327.5	25.0	211	2 Q8YDU7	Q8YDU7 mus musculus
6	326	24.9	212	1 SCTM_MOUSE	Q9J159 mus musculus
7	325	24.8	212	2 Q9D966	Q9D966 mus musculus
8	324	24.8	212	2 Q8CE38	Q8CE38 mus musculus
9	103.5	7.9	550	2 Q52289	Q52289 pseudomonas
10	102.5	7.8	506	1 SHS1_BOVIN	O45631 bos taurus
11	101	7.7	334	2 Q8XZ99	Q8XZ99 ralstonia s
12	98.5	7.5	304	2 Q8BPN5	Q8BPN5 caenorhabdi
13	98	7.5	542	2 Q8NHNS	Q8NHNS homo sapien
14	97	7.4	894	2 Q9V5F7	Q9V5F7 drosophila
15	96.5	7.4	550	2 Q8BP09	Q8BP09 pseudomonas
16	96	7.3	442	2 Q8HB33	Q8HB33 homo sapien
17	96	7.3	917	2 Q6NNX3	Q6NNX3 drosophila
18	96	7.3	969	2 Q96KG6	Q96KG6 homo sapien
19	95	7.3	1140	2 Q8OT91	Q8OT91 mus musculus
20	94.5	7.2	393	2 Q82004	Q82004 lycopersico
21	94.5	7.2	690	2 Q62AS3	Q62AS3 burkholderi
22	94.5	7.2	690	2 Q63JU4	Q63JU4 burkholderi
23	91.5	7.0	864	2 Q6P779	Q6P779 rattus norv
24	91.5	7.0	1260	2 Q7NR05	Q7NR05 chromobacte
25	91	7.0	6620	2 Q96AA2	Q96AA2 homo sapien
26	90.5	6.9	271	2 Q9S218	Q9S218 streptomyce
27	90	6.9	1088	2 Q9C9D7	Q9C9D7 arabidopsis
28	89	6.8	798	2 Q8YVJ4	Q8YVJ4 anabaena sp
29	88.5	6.8	352	2 Q63NY8	Q63NY8 burkholderi
30	88	6.7	327	2 Q8RIG9	Q8RIG9 oryza sativ
31	87.5	6.7	298	2 Q804R4	Q804R4 brachydanio

32	87.5	6.7	372	2	Q90Y50	Q90Y50 brachydanio
33	87.5	6.7	645	2	Q8MJZ5	Q8MJZ5 pan troglod
34	87.5	6.7	947	2	Q8BKK7	Q8BKK7 mus musculus
35	87	6.7	259	2	Q82SL8	Q82SL8 nitrosomona
36	86.5	6.6	395	2	Q88LS8	Q88LS8 pseudomonas
37	86.5	6.6	420	2	Q87DX4	Q87DX4 xyella fae
38	86.5	6.6	930	2	Q02472	Q02472 theiller's e
39	86	6.6	493	1	MOCR_RHIME	P49309 rhizobium m
40	86	6.6	508	2	Q80ZD5	Q80ZD5 rattus norv
41	85.5	6.5	407	2	Q9D2J4	Q9D2J4 mus musculus
42	85.5	6.5	418	2	Q8MI84	Q8MI84 pongo pygma
43	85.5	6.5	420	2	Q9PDS5	Q9PDS5 xyella fae
44	85.5	6.5	430	2	Q8MI85	Q8MI85 pongo pygma
45	85.5	6.5	922	2	Q88495	Q88495 theiller's e

ALIGNMENTS

RESULT 1
SCTM_HUMAN
ID SCTM_HUMAN STANDARD; PRT; 248 AA.
AC Q8WV66: Q00466;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Secreted and transmembrane protein 1 precursor (Protein K12).
GN Name=SECTM1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.
RX MEDLINE=98149980; PubMed=9480746; DOI=10.1006/geno.1997.5151;
RA Sientz-Kesler K.A., Hale L.P., Kaufman R.E.;
RT "Identification and characterization of K12 (SECTM1), a novel human
RT gene that encodes a Golgi-associated protein with transmembrane and
RT secreted isoforms.";
RN Genomics 47:327-340(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUES=Liver;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.P., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Lequellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP INTERACTION WITH CD7.
RX MEDLINE=20119303; PubMed=10652336; DOI=10.1074/jbc.275.5.3431;
RA Lyman S.D., Escobar S., Rousseau A.-M., Armstrong A., Fanslow W.C.;
RT "Identification of CD7 as a cognate of the human K12 (SECTM1)
RT protein.";
RL J. Biol. Chem. 275:3431-3437(2000).
CC -!- SUBUNIT: Interacts with CD7.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Probable). Also

CC found as secreted.
CC TISSUE SPECIFICITY: Detected at the highest levels in peripheral
CC blood leukocytes and breast cancer cell lines. Found in leukocytes
CC of the myeloid lineage, with the strongest expression observed in
CC granulocytes and no detectable expression in lymphocytes.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; U77643; AAC52044.1; --
CC EMBL; BC017716; AAH17716.1; --
CC Genew; HGNC:10707; SBC7M1.
CC H-InvDB; HIX0014254; --
CC MIM; 602602; --
CC InterPro: IPR007110; Ig-like.
KW Signal; Transmembrane
FT SIGNAL 1 Potential.
FT CHAIN 28 Secreted and transmembrane protein 1.
FT DOMAIN 29 248 Extracellular (Potential).
FT TRANSMEM 145 166 Potential.
FT TRANSMEM 146 166 Cytoplasmic (Potential).
FT DOMAIN 167 248 Potential.
FT DISULFID 38 55 Potential.
FT CARBOHYD 56 56 N-linked (GlcNAc...) (Potential).
FT CONFLICT 191 191 V -> F (in Ref. 2).
SQ SEQUENCE 248 AA; 27039 MW; 21E3065B67920487 CRC64;
Query Match 100.0%; Score 1308; DB 1; Length 248;
Best Local Similarity 100.0%; Pred. No. 1.4e-107;
Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MQTCPLAPGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSGENTVMSCNISNA 60
DB 1 MQTCPLAPGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSGENTVMSCNISNA 60
QY 61 FSHVNIKLRAHQSSAIFNEVAPGYFSRSDGWLQVQGGVAQLVKGARDSHAGLYMHLV 120
DB 61 FSHVNIKLRAHQSSAIFNEVAPGYFSRSDGWLQVQGGVAQLVKGARDSHAGLYMHLV 120
QY 121 GHORNNRQVTLVSGAEPSQAPDTGFWPVPVAVTAVFLLVALVMPFAYRCRCSQORREK 180
DB 121 GHORNNRQVTLVSGAEPSQAPDTGFWPVPVAVTAVFLLVALVMPFAYRCRCSQORREK 180
QY 181 KFFLELPQMKVAALRAGAQQLSRASAEELWTPDSEPTPRPLALVFKPSPLGALELLSPQ 240
DB 181 KFFLELPQMKVAALRAGAQQLSRASAEELWTPDSEPTPRPLALVFKPSPLGALELLSPQ 240
QY 241 LFPYAADP 248
DB 241 LFPYAADP 248
RESULT 2
Q921W8 PRELIMINARY; PRT; 192 AA.
ID Q921W8
AC Q921W8; 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE CDNA sequence BC010462.
GN Name=BC010462;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Mammary tumor;
RC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RX

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Spletton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udlin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Besak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Villalon D.K., Muny D.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Richards S., Worley K.C., Sodergren E.J., Lu X., Gibbs R.A.,
RA Whiting M., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
RA and mouse cDNA sequences."
EL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Mammary tumor;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC010462; AAH10462.1; --
DR MGI; MGI:2384805; BC010462.
DR SMART; SMO0409; IG; 1.
SQ SEQUENCE 192 AA; 21438 MW; 13D8989D3F07AE1F CRC64;
Query Match 29.2%; Score 382; DB 2; Length 192;
Best Local Similarity 43.6%; Pred. No. 9.3e-26;
Matches 85; Conservative 31; Mismatches 55; Indels 24; Gaps 6;
QY 1 MQTCPLAPGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSGENTVMSCNISNA 60
DB 1 MQTCPLAPGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSGENTVMSCNISNA 59
QY 61 FSHVNIKLRAHQSSAIFNEVAPGYFSRSDGWLQVQGGVAQLVKGARDSHAGLYMHLV 120
DB 60 FTDVTIQLSANGKDTIFDKKPGQNFWSRQVQGGVAQLVKGARDSHAGLYMHLV 119
QY 121 GHORNNRQVTLVSGAEPSQAPDTGFWPVPVAVTAVFLLVALVMPFAYRCRCSQORREK 160
DB 120 GHORNNRQVTLVSGAEPSQAPDTGFWPVPVAVTAVFLLVALVMPFAYRCRCSQORREK 177
QY 161 VA-LVMPFAYRCRCS 174
DB 178 VAGFTTFIYRHRSS 192
RESULT 3
Q6P781 PRELIMINARY; PRT; 218 AA.
ID Q6P781
AC Q6P781; 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Hypothetical protein MGC72571.
GN Name=MGC72571;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Prostate;
EX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA	Fahcy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA	Krzywinski M.I., Skalska U., Smalusz D.E., Scherch A., Schein J.E.,
RA	Jones S.J., Marra M.A.,
RT	"generation and initial analysis of more than 15,000 full-length human
RT	and mouse cDNA sequences.";
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN	[2]
RP	SEQUENCE FROM N.A.
RC	TISSUE=Kidney;
RA	Director MGC Project;
RL	Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
DR	EMBL; BC078937; AAH78937.1; -
DR	InterPro; IPR007110; Ig-like.
KW	Hypothetical protein.
SQ	SEQUENCE 201 AA; 22762 MW; 790B959BA4839B98 CRC64;
Query Match 27.4%; Score 358; DB 2; Length 201;	
Best Local Similarity 41.8%; Pred. No. 1.3e-23;	
Matches 79; Conservative 25; Mismatches 61; Indels 24; Gaps	
Qy	16 LGTLLFLAAISLAONEGWDSPTCTGCVVSVSGENTVMSCNISAFSHWNKLRHGQS 75
Dd	::: ::: ::: ::: ::: ::: ::: ::: ::: :::
Dd	15 LFSILLMVSLNAQKSWDRPCTENIVSPGRPAVMSCNISNTFTDVTIQLTAHGKDR 74
Qy	76 AINEVAPGVFSDRGWLOVOGGVAOLVKARGASHGLYMWHLVGHORNNHQVTLVSG 135
Dd	::: ::: ::: ::: ::: ::: ::: ::: :::
Dd	75 TIFKKKPGGNFWSGWEVLHVQGGQAQLVKDAQDDHRGIYLQWLHGRGRRYYRNTILNV- 133
Qy	136 AEPO-----SAPD-----TGFWPVAVTVATFVLIVA-----LVNFAYVRCCR 173
Dd	::: ::: ::: ::: ::: ::: ::: ::: :::
Dd	134 -EESNDKVYDTLRFSLPDQVKSQLEAKGTLMGVLIANLGVPGILALICYRHCS 192
Qy	174 SQQRREKFF 182
Dd	:: ::
Dd	193 QKPHWVRRF 201
RESULT 5	
Q8VDU7	ID QBVDU7 PRELIMINARY; PRT; 211 AA.
AC	QBVDU7
DT	01-MAR-2002 (TrEMBLrel. 20, Created)
DT	01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE	Setml protein.
GN	Name=Setml;
OS	Mus musculus (Mouse).
OC	Eukaryota; Metazoa;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX	NCBI_TaxId=10090;
RX	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=CZECH II; TISSUE=Mammary tumor;
RX	MEDLINE=2238257; Pubmed=12477932; DOI=10.1073/pnas.242603899;
RA	Austrberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA	Klausner R.D., Collins F.S., Wagner H., Shenmen C.M., Schuler G.D.,
RA	Altschul S.F., Zeeberg B., Buote K.H., Schaefer C.F., Bhat N.K.,
RA	Hopkins R.F., Jordan H., Max S.I., Wang J., Hsieh F.,
RA	Datchenko L., Marushima K., Farmer A.A., Rubin G.M., Hong L.,
RA	Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA	Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullány S.J.,
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA	Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA	Fahcy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA	Krzywinski M.I., Skalska U., Smalusz D.E., Scherch A., Schein J.E.,

RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CZECH II; TISSUE=Mammary tumor;
 RA Strausberg R.;
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC020159; AAH20159.1; -;
 DR MGD; MGI:1929083; Sectm1.
 DR SMART; SM00409; IG; 1.
 SQ SEQUENCE 211 AA; 23433 MW; 9E6FB436592B2D8 CRC64;
 Query Match 25.0%; Score 327.5; DB 2; Length 211;
 Best Local Similarity 40.0%; Pred. No. 6.9e-21;
 Matches 78; Conservative 24; Mismatches 70; Indels 23; Gaps 3;
 QY 10 GHVSQALGTLFLAASLSAONEGWDSPTCTEGVSVSGENTVMSCNISNAPSHVNIKL 69
 Db 10 GLFPMMLWALLLLAASLNAYHVDKPCCTEHEVSVNRGSRVVMACNISNLRDVTIELV 69
 QY 70 AHQGESALFNEVAPGYFSDRGVQLVQGGVQLVKGARDSHAGLYMHLVGHQNRNQV 129
 Db 70 TSEKTSIIIFNTPPGNYSKDSWQLHIQGVQALVITDAQCKSGEYWKLRGLQAEFKNF 129
 QY 130 TLVSGAEPQAPDTGFWPVP-----AVVTAVFILLVA-----LVMPFWTRCRCSQ 177
 Db 130 NLIVNAADRKQTEFLPVTKVPDPTAVMTVEIIIAITTTIITIGVFWVY----- 181
 QY 178 REKKFFLEPQMKVA 192
 Db 182 ---KQFPVAPQIQMS 193
 RESULT 6
 SCTM MOUSE STANDARD; PRT; 212 AA.
 ID SCTM MOUSE
 AC Q9JL59;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Secreted and transmembrane protein 1 precursor.
 GN Name=Secretm;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A., AND INTERACTION WITH CD7.
 RC TISSUE=Colon;
 RX MEDLINE=201119303; PubMed=106523236; DOI=10.1074/jbc.275.5.3431;
 RA Lyman S.D., Escobar S., Rousseau A.-M., Armstrong A., Fanslow W.C.;
 RT "Identification of CD7 as a cognate of the human K12 (SECTM1)
 protein.";
 RL J. Biol. Chem. 275:3431-3437 (2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ussid T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Musny D.M., Sodergren B.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Ketterman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 CC -I- SUBUNIT: Interacts with CD7.
 CC -I- SUBCELLULAR LOCATION: Type I membrane protein (Probable). Also
 found as secreted (By similarity).
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF210700; AAF30406.1; -;
 CC EMBL; BC010805; AAH10805.1; -;
 CC MGD; MGI:1929083; Sectm1.
 DR InterPro; IPR003599; IG-like.
 DR SMART; SM00409; IG; 1.
 KW Signal; Transmembrane.
 FT SIGNAL 1 28
 FT CHAIN 29 212
 FT DOMAIN 29 160
 FT TRANSMEM 161 181
 FT DOMAIN 182 212
 FT DISULFID 38 55
 FT CARBOHYD 56 56
 FT CARBOHYD 85 85
 FT CARBOHYD 114 114
 FT CARBOHYD 130 130
 FT SEQUENCE 212 AA; 23477 MW; 75113E877A2C5B87 CRC64;
 Query Match 24.9%; Score 326; DB 1; Length 212;
 Best Local Similarity 37.8%; Pred. No. 9.4e-21;
 Matches 74; Conservative 25; Mismatches 73; Indels 24; Gaps 2;
 QY 10 GHVSQALGTLFLAASLSAONEGWDSPTCTEGVSVSGENTVMSCNISNAPSHVNIKL 69
 Db 10 GLFPMMLWALLLLAASLNAYHVDKPCCTEHEVSVNRGSRVVMACNISNLRDVTIELV 69
 QY 70 AHQGESALFNEVAPGYFSDRGVQLVQGGVQLVKGARDSHAGLYMHLVGHQNRNQV 129
 Db 70 TSEKTSIIIFNTPPGNYSKDSWQLHIQGVQALVITDAQCKSGEYWKLRGLQAEFKNF 129
 QY 130 TLVSGAEPQAPDTGFWPVP-----AVVTAVFILLVA-----LVMPFWTRCRCSQ 176
 Db 130 NLIVNAADRKQTEFLPVTKVPDPTAVMTVEIIIAITTTIITIGVFWVY----- 182
 QY 177 REKKFFLEPQMKVA 192
 Db 183 ---KQFPVAPQIQMS 194
 RESULT 7
 Q9D966 PRELIMINARY; PRT; 212 AA.
 ID Q9D966
 AC Q9D966
 DT 01-JUN-2001 (TEMBLrel. 17, Created)
 DT 01-JUN-2001 (TEMBLrel. 17, Last sequence update)
 DT 01-MAR-2004 (TEMBLrel. 26, Last annotation update)
 DE Mus musculus 10 day old male pancreas cDNA, RIKEN full-length enriched
 library, clone:1810003C24 product:secreted and transmembrane 1, full
 insert sequence.
 GN Name=Secretm;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

```

OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Pancreas;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44 (1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Pancreas;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690 (2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Pancreas;
RA The FANTOM Consortium;
RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Pancreas;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M., Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630 (2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Pancreas;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P., Konno H., Akiyama J., Nishi K., Kitsuunai T., Tashiro H., Itoh M., Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A., Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K., Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M., Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J., Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771 (2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Pancreas;
RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H., Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M., Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F., Imokani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H., Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C., Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M., Okazaki Y., Okido T., Owa C., Saito K., Saito R., Sakai C., Sakai K., Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T., Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T., Teijima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK007321; BAB24956.1; -
DR MGD; MGI:1929083; Sectm1.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR SMART; SM00409; IG; 1.
XW Transmembrane.
SQ SEQUENCE 212 AA; 23559 MW; C8631ABFC724FE9E CRC64;
Query Match 24.8%; Score 325; DB 2; Length 212;
Best Local Similarity 37.8%; Pred. No. 1.2e-20;
Matches 74; Conservative 25; Mismatches 73; Indels 24; Gaps 2;

```

RESULT 8

Q8CE38

ID Q8CE38 PRELIMINARY; PRT; 212 AA.

AC Q8CE38;

DT 01-MAR-2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE Mus musculus 10 days neonate skin cDNA, RIKEN full-length enriched DE library, clone:4732489J23 product:secreted and transmembrane 1, full insert sequence.

GN Name=Secret1;

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Skin;

RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;

RA Carninci P., Hayashizaki Y.;

RT "High-efficiency full-length cDNA cloning.";

RL Meth. Enzymol. 303:19-44 (1999).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Skin;

RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;

RA RIKEN FANTOM Consortium;

RT "Functional annotation of a full-length mouse cDNA collection.";

RL Nature 409:685-690 (2001).

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Skin;

RA The FANTOM Consortium;

RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs.";

RL Nature 420:563-573 (2002).

RN [4]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Skin;

RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;

RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M., Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;

RT "Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes.";

RL Genome Res. 10:1617-1630 (2000).

RN [5]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Skin;

RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;

RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P., Konno H., Akiyama J., Nishi K., Kitsuunai T., Tashiro H., Itoh M., Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A., Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K., Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,

Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J., Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.; "RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer."; Genome Res. 10:1757-1771(2000).

[6]

SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Skin;

RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P., Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W., Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T., Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T., Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S., Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M., Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y., Saito K., Saitho H., Sakai C., Sakai K., Sakazume N., Sano H., Sasaki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M., Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T., Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.; Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.

RL ENBL; AK029082; BAC26284.1; --

DR MGD; MGI:1929083; Sectml.

DR GO; GO:0016021; C:integral to membrane; IEA.

DR InterPro; IPR003599; IG.

DR InterPro; IPR007110; IG-like.

DR SMART; SM00409; IG; 1.

KW Transmembrane.

SQ SEQUENCE 212 AA; 23559 MW; C863064DCE36EB7E CRC64;

Query Match 24.8%; Score 324; DB 2; Length 212;

Best Local Similarity 37.8%; Pred. No. 1.4e-20;

Matches 74; Conservative 25; Mismatches 73; Indels 24; Gaps 2;

QY 10 GHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSNINAFSHVNIKL 69

DB 10 GLFPRMLWALLLLAASLNAYNHVWDKPCCTEHEVSVNRGSRVWACNINLRDVTIELV 69

QY 70 AHGQESAIENFVARGYFSDGQLOVGGVAQLVIKARDSHAGLVYHVLGHQRNROV 129

DB 70 TSKTSTIIFNKTPGNTSKDSWQLHIQGGQALVITDAQKHGEVYKWLGRGFAEFKNF 129

QY 130 TLEVSGAEPOSAPDTGFWFV-----AVTVAVFILLVALVWFVWYRCRCSQ 176

DB 130 NLIVNAORQKTEDLPTVKFDKPTAVRTEVIIIAITIIITIGVFWY----- 182

QY 177 RREKKFFLLEPQMKVA 192

DB 183 ----KQPPVAFQIQMS 194

RESULT 9

Q52289 PRELIMINARY; PRT; 550 AA.

AC Q52289

DT 01-NOV-1996 (TrEMBLrel. 01, Created)

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE UxpA protein.

GN Name=uxpA;

OS Pseudomonas putida.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;

OC Pseudomonadaceae; Pseudomonas.

OC NCBI_TaxID=303;

OX [1]

RN SEQUENCE FROM N.A.

RC STRAIN=WCS358;

RA de Groot A., Krijger J.J., Filloux A., Tommassen J.; "Characterization of type II protein secretion (xcp) genes in the plant growth-stimulating Pseudomonas putida, strain WCS358."; Mol. Gen. Genet. 250:491-504(1996).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=WCS358;

RA de Groot A.; Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.

RL ENBL; X81085; CAAS6977.1; --

DR PIR; S64725; S64725;

SQ SEQUENCE 550 AA; 59118 MW; 684D79D0D279868B CRC64;

Query Match 7.9%; Score 103.5; DB 2; Length 550;

Best Local Similarity 23.2%; Pred. No. 1.4;

Matches 72; Conservative 42; Mismatches 118; Indels 79; Gaps 16;

QY 2 QTCPLAPPGHVSOA-IGTLFLAA-----SLSAQN-EGWDSPICTEGVSVS 46

DB 90 QLAPLLDAGRADEANLGGYAVLGLLEQLRGEAGAGNSLTLENGQGWG-----SGLAYIT 145

QY 47 WGENVTWSCNISAFSHVNIKLRA--HGQESAIENFVARGYFSDR-----DGWQ 92

DB 146 QGESGVGQSOLLGSEARVSDERVLFPQRSAALYRQASATTLGAGLADEQRQALGLEPLQ 205

QY 93 LOVGQGVAAQLVIK-----ARDSHAGLVYHVLGHQRNROVTLVSGAEPOSAP-- 142

DB 206 LFERGG-ARIAVGVTDVPAQDKASLKQWYQALLPVFQQRREADLVVALADVGTGPGL 264

QY 143 -----DTGFWFVPAVTVAVFILLVALVWFV-----YRCRC-----SQ 176

DB 265 WLAERLPAIDLLLCARGODLWPTPVQATQASGRVP-VLPAGCGSGAFRLRCQVAGOW 323

QY 177 RREKKFF-----LLEPQMKVAALRAGAQGLSRASAEIWTDPSEPTPR-PLALVFKPSPL 230

DB 324 QFEGRFPTFEQTLSPAAQVRAQGLQALNQQRAGHAWL--DOPLARAPQALWRRDTRG 381

231 GALELLSPQL 241

382 GSWDRLLHQAL 392

RESULT 10

SHS1_BOVIN

ID SHS1_BOVIN STANDARD; PRT; 506 AA.

AC Q46631; Q46632;

DT 10-OCT-2003 (Rel. 42, Created)

DT 10-OCT-2003 (Rel. 42, Last sequence update)

DT 05-JUL-2004 (Rel. 44, Last annotation update)

DE Protein-tyrosine phosphatase non-receptor type substrate 1 precursor (SHP substrate-1) (SHPs-1) (Inhibitory receptor SHPS-1) (Signal-regulatory protein alpha-1) (SIRP-alpha-1) (MyD-1 antigen).

DE Name=PTPNS1; Synonyms=MYD1, SHPS1, SIRP;

OS Bos taurus (Bovine).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

OC Bovinae; Bos.

OC NCBI_TaxID=9913;

OX [1]

RN SEQUENCE FROM N.A.; AND VARIANTS SER-23; ALA-28; LEU-61; ARG-70;

RP HTS-120; 125-GLN; 129-GLY; 129-HIS; 132-VAL; ASN-145; VAL-153;

RP ASP-203; ARG-261; LEU-302; LEU-316; ARG-337; ASN-367; LEU-422; PHE-429

AND GLU-433.

RC STRAIN=Friesian; TISSUE=Peripheral blood;

RX MEDLINE=98143722; PubMed=9485180;

RX DOI=10.1002/(SICI)1521-4141(199801)28:01<1::AID-IMMU1>3.3.CO;2-M;

RA Brooke G.P., Parsons K.R., Howard C.J.;

RT "Cloning of two members of the SIRP alpha family of protein tyrosine phosphatase binding proteins in cattle that are expressed on monocytes and a subpopulation of dendritic cells and which mediate binding to CD4 T cells.;"

RT Eur. J. Immunol. 28:1-11(1998).

CC -!- FUNCTION: Immunoglobulin-like cell surface receptor for CD47. Acts as docking protein and induces translocation of PTPN6, PTPN11 and other binding partners from the cytosol to the plasma membrane. Supports adhesion of cerebellar neurons, neurite outgrowth and glial cell attachment. May play a key role in intracellular signaling during synaptogenesis and in synaptic function. Involved in the negative regulation of receptor tyrosine kinase-coupled

cellular responses induced by cell adhesion, growth factors or insulin. Mediates negative regulation of phagocytosis, mast cell activation and dendritic cell activation. CD47 binding prevents maturation of immature dendritic cells and inhibits cytokine production by mature dendritic cells (By similarity).

-!- SUBUNIT: Binds PRP11 when tyrosine-phosphorylated, except in macrophages, where it primarily binds PTPN6. Binds GRB2 in vitro. Binds JAK2 irrespective of its phosphorylation status and forms a stable complex. Binds SCAP1 and/or SCAP2. The resulting complex recruits FYB. Binds FGR and PTK2B (By similarity).

-!- SUBCELLULAR LOCATION: Type I membrane protein.

-!- TISSUE SPECIFICITY: Highly expressed in spleen macrophages. Detected in skin dendritic cells.

-!- PTM: Phosphorylated on tyrosine residues (By similarity).

-!- SIMILARITY: Contains 2 immunoglobulin-like C1-type domains.

-!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; Y11045; CAA71942.1; -
 EMBL; Y11046; CAA71943.1; -
 HSP; O88792; I197.
 InterPro; IPR007110; Ig-like.
 InterPro; IPR003597; Ig_C1.
 InterPro; IPR003006; Ig_MHC.
 Pfam; PF00047; Ig_3.
 SMART; SM00407; IG_C1; 2.
 PROSITE; PS50835; IG_LIKE; 3.
 PROSITE; PS00290; IG_MHC; FALSE NEG.
 Glycoprotein; Immunoglobulin domain; Phosphorylation; Polymorphism;
 Repeat; SH3-binding; Signal; Transmembrane.
 SIGNAL 1 29
 CHAIN 30 506
 DOMAIN 30 371
 TRANSMEM 372 392
 DOMAIN 393 506
 DOMAIN 30 145
 DOMAIN 148 248
 DOMAIN 255 348
 DISULFID 55 121
 DISULFID 170 228
 SITE 273 331
 SITE 432 435
 SITE 441 446
 SITE 455 458
 SITE 472 475
 SITE 498 501
 MOD_RES 431 431
 MOD_RES 455 455
 MOD_RES 472 472
 MOD_RES 498 498
 CARBOHYD 92 92
 CARBOHYD 167 167
 CARBOHYD 179 179
 CARBOHYD 204 204
 CARBOHYD 210 210
 CARBOHYD 246 246
 CARBOHYD 270 270
 CARBOHYD 292 292
 CARBOHYD 311 311
 CARBOHYD 319 319
 CARBOHYD 344 344

FT CARBOHYD 365 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 368 N-linked (GlcNAc...) (Potential).
 FT VARIANT 23 N -> S.
 FT VARIANT 28 T -> A.
 FT VARIANT 61 S -> L.
 FT VARIANT 70 K -> R.
 FT VARIANT 120 Y -> H.
 FT VARIANT 125 R -> Q.
 FT VARIANT 127 E -> G.
 FT VARIANT 129 R -> H.
 FT VARIANT 132 M -> V.
 FT VARIANT 145 S -> N.
 FT VARIANT 153 L -> V.
 FT VARIANT 203 N -> D.
 FT VARIANT 261 G -> R.
 FT VARIANT 302 F -> L.
 FT VARIANT 316 F -> L.
 FT VARIANT 337 G -> R.
 FT VARIANT 367 S -> N.
 FT VARIANT 422 Q -> L.
 FT VARIANT 429 I -> F.
 FT VARIANT 433 D -> E.
 SQ SEQUENCE 506 AA; 55093 MW; 6B7E310677FCE9CB CRC64;
 Query Match 7.8%; Score 102.5; DB 1; Length 506;
 Best Local Similarity 27.0%; Pred. NO. 1.5;
 Matches 51; Conservative 25; Mismatches 82; Indels 31; Gaps 7;
 QY 4 CPLAPPCHVSQALCTLLFLAASLSAQNEGWSPTCTEGVSVSWGNTVMSCNISNAP-S 62
 DB 228 CEVA---HVTIQGGPPLRGNTANLS---ETIRVPPTLEITGSPSAGNVNVTQVKNKPYPR 281
 QY 63 HVNIKLRAGH-----QESAIFNEVAPGVFSRDGWLQVQGGVAQLVIKGRDASHAGLYNM 117
 DB 282 HLQLTWLENGNMSRTEAASVFENKDTGFNQTSLVFNSSAHEAVLTCQVEHD----- 336
 QY 118 HLVGHQRNRRQVTLVSGAEPQAPDTGFWPVP-----AVTAVPILVALVWFAM 168
 DB 337 ---GQPAVSKNHTLEVSA--PQKODTGTGTPGNDNNWTSIFIVGVVWCALLVALLIAL 391
 QY 169 YRCRCSQOR 177
 DB 392 YLLRIRQNK 400
 RESULT 11
 Q8XZ99 PRELIMINARY; PRT; 334 AA.
 AC Q8XZ99;
 DT 01-MAR-2002 (Tremblrel. 20, Created)
 DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
 DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
 DE PROBABLE ALCOHOL DEHYDROGENASE-LIKE OXIDOREDUCTASE PROTEIN (EC 1.1.1.-)
 DE).
 GN Name=RS03800; OrderedLocusNames=RS03800;
 OS Ralstonia solanacearum (Pseudomonas solanacearum).
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Burkholderiaceae; Ralstonia.
 OX NCBI_TaxID=305;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GMT1000;
 RX MEDLINE=21681879; PubMed=11283852; DOI=10.1038/415497a;
 RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
 RA Ariat M., Billault A., Brottier P., Camus J.C., Catolico L.,
 RA Chander M., Choise N., Claudel-Renard C., Cunnac S., Demange N.,
 RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,
 RA Siguer P., Thebault P., Whalen M., Wincker P., Levy M.,
 RA Weissenbach J., Boucher C.A.;
 FT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
 RL Nature 415:497-502(2002).
 CC -!- COPACTOR: Zinc (By similarity).
 CC -!- SIMILARITY: Belongs to the zinc-containing alcohol dehydrogenase

```

CC family.
DR EMBL: AL646064; CAD15207.1; -.
DR HSSP; P75691; 1UUF
DR GO: GO:0004024; F:alcohol dehydrogenase activity, zinc-dependent; IEA.
DR GO: GO:0016491; F:oxidoreductase activity; IEA.
DR GO: GO:0008270; F:zinc ion binding; IEA.
DR Pfam; PF00107; ADH zinc N; 1.
DR PROSITE; PS00059; ADH ZINC; UNKNOWN 1.
DR Complete proteome; Meta1-binding; Oxidoreductase; Zinc.
KW SEQUENCE 334 AA; 35204 MW; 0F0176E32E8F3BA2 CRC64;

Query Match 7.7%; Score 101; DB 2; Length 334;
Best Local Similarity 24.6%; Pred. No. 1.3;
Matches 67; Conservative 28; Mismatches 83; Indels 94; Gaps 16;

QY 5 PLAFPHVSOALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSNCISNAFHV 64
DB 54 PALIPGH--EIVGVTDACGPGVSLTGE-----RVGVPWLGSCTGCPFCR----- 98
QY 65 NIKLRAHQESAIENEVAFGYFSRDGWLQVOGGVAQLVIKAR-----DSHAG- 113
DB 99 -----RDQENLCDRPFTGY-TRD-----GGVAEYTVCDARYCLPIPARYDDAHAAP 144
QY 114 LYMHLVGHQRRNRQVTLVSGAEPOSAPDTGFWPVPVAVTAVFILLVALVMFAWYRCRC 173
DB 145 LLCAGLIGYR-----TLRMAG-DARRLGIYFGAAHLVLTQIAV----- 182
QY 174 SQORREKFFLLPQMKVAALRAGAACQGLSRASAEIWTDPSE-PTPRPL--ALVFKPQPL 230
DB 183 AEQREVAF-----TRAGDTGAQO--LAROTGACWAGSEAPPEAPLDAALIF--APV 231
QY 231 GALELLSPQ-----PLFFY 244
DB 232 GALVPLALQAVKVGTVCGGIHMSDIPAPPY 263

RESULT 12
Q9BPN5 PRELIMINARY; PRT; 304 AA.
AC Q9BPN5;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein Y32G9A.8.
GN Name=Y32G9A.8; ORFNames=Y32G9A.8;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RG WormBase Consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
RL investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Bemis G., Courtney L.;
RT "The sequence of C. elegans cosmid Y32G9A.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RG WormBase Consortium;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.

```

```

DR EMBL: AC087794; AAG53701.2; -.
DR WormBase; WEGene00021305; Y32G9A.8.
DR WormPep; Y32G9A.8; CE31101.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig_c2.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00408; IGC2; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
KW Hypothetical protein_IG_LIKE.
KW SEQUENCE 304 AA; 34112 MW; AEF48BB42A0DB012 CRC64;

Query Match 7.5%; Score 98.5; DB 2; Length 304;
Best Local Similarity 22.1%; Pred. No. 1.9;
Matches 36; Conservative 28; Mismatches 52; Indels 47; Gaps 6;

QY 49 ENTVMSCNISNAESHVNIK-----LRAHQESAIENEVAPGYFS----- 87
DB 121 DNSIIIGNMF-AYSHVPVKKNNETWELKSESEFTVGVPAVAPLDSMARIQCPVGYPE 179
QY 88 -----RDGWLQVOGGV-----AQLVIKARGDSHAGLYMMHLVGHQRRNRQVTLVSGAE 137
DB 180 PQIWIYKDKPLEIGRVKFTAGVLSIEGAQEDAGVYRCEA-----TNQFFVQIDGPE 233
QY 138 POSA-----PDTGFWPVPVAVTAVFILLVALVMFAWYRC 171
DB 234 QHFAVKLDQELRLIGDSYGMPLAIIILILLFLVFTQCRC 276

RESULT 13
Q8NHNS PRELIMINARY; PRT; 542 AA.
ID Q8NHNS;
AC Q8NHNS;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Obscurin (fragment).
GN Name=OBSCN;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Young P., Ehler E., Gautel M.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Gautel M.S.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ314903; CAC85750.1; -.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig_c2.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00408; IGC2; 1.
DR PROSITE; PS50835; IG_LIKE; 6.
DR NON_TER 1 542
FT NON_TER 542
SQ SEQUENCE 542 AA; 58725 MW; F2F41C47CCF4F198 CRC64;

Query Match 7.5%; Score 98; DB 2; Length 542;
Best Local Similarity 23.2%; Pred. No. 4.1;
Matches 60; Conservative 35; Mismatches 94; Indels 70; Gaps 11;

QY 11 HVSQALGTLFLAASLSAQNEGWDSPICTEGV-----VSVWGENTVMSNCISNA 60
DB 248 HGAQVLDSAIY---SCRVGAEGQDPFVQVEEVAAKFCRLLEPVCGELGTVTIACELSPA 304
QY 61 FSHV-----NIKLRHQESAIENEVAPGYFSRDGWLQVOGGVAQLVIKGRDASHAGLY 115
DB 305 CAEVMVRCGNTQLRV-GKR-----FQMAEGPVRSLTVLGLRAEDAGEY 347
QY 116 MMHLVGHQRRNR---QVTLV-----SGAEPOSAPDTGFWPVPVAVTAVFILLVALVM 165

```

```

Db 348 -----VCESRDDHTSQAQLTVSVPRVVKFMGSLSTVVAEEGGEATFOCVSPSDVAVV----- 399
QY 166 FAWYRCRCQORREKKFFLLBQMKVAALRAGAAGQGLSRASAE-----WTPDSBPTPRP 220
Db 400 -----LLQPSKFAISQSGASHSLTISDLVLEDAQGITVAEGASS 447
QY 221 LAIVFKPSPGLALELLSPQ 239
Db 448 AALRVREAPVLFKKLEPQ 466

RESULT 14
Q9V5F7
ID Q9V5F7 PRELIMINARY; PRT; 894 AA.
AC Q9V5F7;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE CG2292-PA.
GN ORFNames=CG2292;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
SEQUENCE FROM N.A.
RP MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RX Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abril J.F., Aghayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fouts L.C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobery C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskae R. Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wasserman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodage, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhou S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
RN [2]
SEQUENCE FROM N.A.
RX MEDLINE=22426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R.,
RA Paciel J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,

```

```

RA Svirskae R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun: Release 3 of the Drosophila
melanogaster euchromatic genome sequence."
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN [3]
SEQUENCE FROM N.A.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Krommiller B., Carlson J., Svirskae R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celniker S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
a genomics perspective."
RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
RN [4]
SEQUENCE FROM N.A.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
systematic review."
RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
RN [5]
SEQUENCE FROM N.A.
RP FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN [6]
SEQUENCE FROM N.A.
RG FlyBase;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE039831; AAF58853.2; -.
DR FlyBase; FGN0033479; CG2292.
DR InterPro; IPR007070; PIGN.
DR Pfam; PF04987; PIGN; 1.
SQ SEQUENCE 894 AA; 101828 MW; 9BD68206E9595187 CRC64;

Query Match 7.4%; Score 97; DB 2; Length 894;
Best Local Similarity 22.7%; Pred. No. 8.8;
Matches 62; Conservative 33; Mismatches 88; Indels 90; Gaps 15;

QY 9 PGHVSQAQGLTLLFLAASLAQNEGWDS-PICTEGVVSVS-----NGENTVMSCNISNAPS 62
Db 95 PGHITLIAG-----LYEDPSAVLRGWKSNPIDPDTVFNRSQTYANGANDVL-----NVFS 145
QY 63 HVN-----IKLRAHQSSAIFNEVAPGY--FSRDGWLQVQGGVAQLVIKGARDSHAGLYM 116
Db 146 HVSNGGEINLRFYNHDL---DFSPGYDAYEEDWVFKVKLLQKQREALQRAQNVVVF 201
QY 117 WHLVGHQRNQRVTLVSGS-REPOSAPDTGFVFPVAVTAVTAVFILLVALVMPFAYWRCRCQ 175
Db 202 LHLG-----LDTAGHVHKPGAP-----KFRRTL 225
QY 176 ORREK-----KFLLEPQMKVAALRAGAAGQGLSRASAE-----LW-TPDS 214
Db 226 EKTEKGVVAVIQEFERFVPDKRTAYLLT-ADHGMTDSGAHSGSGSPHETDTPPMLMGAGAS 284
QY 215 EPTFRPLALVFKPS-----PLGALELISPOPL 241
Db 285 RAVPKPGCGRTMPNNEGPAFMLELEQAQLTPL 317

RESULT 15
Q88P09
ID Q88P09 PRELIMINARY; PRT; 550 AA.
AC Q88P09;
DT 01-JUN-2003 (TRENBLrel. 24, Created)
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)

```

Sun Mar 20 14:17:43 2005

```

DE Lipoprotein UxpA.
GN Name=uxpA; OrderedLocusNames=PP1044;
OS Pseudomonas putida (strain KT2440).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=160488;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22423060; PubMed=12534463;
RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
RA Brinkac L.M., Beanan M.J., DeBoy R.T., Daugherty S.C., Kolonay J.F.,
RA Madupu R., Nelson W.C., White O., Peterson J.D., Khouri H.M.,
RA Hance I., Chris Lee P., Holtzapple E.K., Scanlan D., Tran K.,
RA Moazzes A., Uterback T.R., Rizzo M., Lee K., Kosack D., Moestl D.,
RA Wedler H., Lauber J., Stjepandic D., Hoheisel J., Straetz M., Heim S.,
RA Kiewitz C., Eisen J.A., Timmis K.N., Duesterhoeft A., Tuemmler B.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative analysis of the
RT metabolically versatile Pseudomonas putida KT2440."
RL Environ. Microbiol. 4:799-808(2002).
DR EMBL; AS016777; AAN66669.1; -.
DR TIGR; PP1044; -.
KW Complete proteome; Lipoprotein.
SQ SEQUENCE 550 AA; 59383 MW; 55750860D219EB2C CRC64;

Query Match 7.4%; Score 96.5; DB 2; Length 550;
Best Local Similarity 22.3%; Pred. No. 5.7;
Matches 67; Conservative 39; Mismatches 96; Indels 99; Gaps 17;

QY 2 QTCPLAPPCHVQA-IGTLLFLAA-----SLSAON-EGWDSPICTEGVSVS 46
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 90 QLAPLIDARQATTAQLGGYAVLAALLEQLRGEAGAGNSITLENGQWNG-----SGLAYLT 145
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 47 WGENTVMSCNISNAFHVNIKRAHQESAI FNEVAPGYF-----SRDGV 91
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 146 QGESGVQGSQ-----LGIEARVSSDERVLPQSTGLYRQASAITLGAGLADAQOQAL 199
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 92 QLQV-----QGGVQLVKG-----ARDSHAGLYMWH---LVGHQNNRQVTLVSGAEPQ 139
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 200 GVQALHVFERRGARIAVVGVTDPYAQDKASLKQWYQSLLPFVEQARHEADLVAVADAG 259
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 140 SAP-----DTGFVFPVAVTVAVFILLVLYMFAW-----YRCRC-- 173
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 260 TGPGLWLAERVPOQIDVLLCARGQDLNPAPEATQASGRYP-VLFAGCRGSGAFRLRCOR 318
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 174 --SQORREKKFF-----LLEP--QMKVAAL-----RAG-----AQGLSRASAEIWT 213
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 319 VAGQWQFDGRFFPAFEQQQLAPAAQLRVAQLQAQLRQORAGHAAWLDOPLARAFQALWRRD 378
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 214 S 214
Db |
QY 379 T 379

```

Search completed: March 20, 2005, 09:49:13
Job time : 100 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 20, 2005, 00:33:52 ; Search time 9530 Seconds
(without alignments)
11084.194 Million cell updates/sec

Title: US-10-080-522-2
Perfect score: 2180
Sequence: 1 ATTCTGCTTCTTTAGCGT.....TTAAATAAGTTCGTCCG 2180

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEnbl.*

- 1: gb_ba.*
- 2: gb_hg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	2180	100.0	2180	6	AR071760 Sequence
2	2180	100.0	2180	6	AR097657 Sequence
3	2180	100.0	2180	6	AR195192 Sequence
4	1951	89.5	2000	6	CQ776411 Sequence
5	1951	89.5	2000	6	CQ861464 Sequence
6	1951	89.5	2000	6	AX055560 Sequence
7	1951	89.5	2000	6	AX333755 Sequence
8	1951	89.5	2000	9	HSU77643 Homo sapien
9	1681.2	77.1	1761	9	BC017716 Homo sapien
10	1230.6	56.4	154252	9	AC132938 Homo sapi
11	1230.6	56.4	171569	9	AC132872 Homo sapi
12	938.6	43.1	1464	6	BD270355 50 human
13	912.8	41.9	1064	6	BD130223 Human sig
14	732.8	33.6	747	12	BT008191 Synthetic
15	732.6	33.6	166458	2	AC084737 Homo sapi
c 16	564.8	25.9	66324	2	AC134680 Homo sapi
17	506.2	23.2	592	6	AX780236 Sequence
c 18	462.4	21.2	71352	2	AC137090 Homo sapi
19	363.8	16.7	401	6	AX778685 Sequence

20	195.4	9.0	1629	10	BC078937	BC078937 Rattus no
21	189	8.7	1582	10	BC010462	BC010462 Mus muscu
22	179	8.2	229839	2	AC094588	AC094588 Rattus no
c 23	179	8.2	249664	2	AC091384	AC091384 Rattus no
c 24	179	8.2	260404	2	AC128909	AC128909 Rattus no
25	174.6	8.0	1249	10	BC061794	BC061794 Rattus no
c 26	170.8	7.8	198011	10	AL662887	AL662887 Mouse DNA
27	159.6	7.3	1182	10	BC020159	BC020159 Mus muscu
28	157	7.2	1076	6	CQ777327	CQ777327 Sequence
29	156.8	7.2	71352	2	AC137090	AC137090 Homo sapi
30	155.4	7.1	1112	6	CQ777326	CQ777326 Sequence
31	155.4	7.1	1112	6	AX055564	AX055564 Sequence
32	155.4	7.1	1112	10	AF210700	AF210700 Mus muscu
33	155.4	7.1	1273	10	BC010805	BC010805 Mus muscu
c 34	139	6.4	73178	2	AC134665	AC134665 Homo sapi
35	77	3.5	125020	9	AF429315	AF429315 Homo sapi
c 36	72.8	3.3	125020	9	AF429315	AF429315 Homo sapi
37	62.4	2.9	101509	2	AC027353	AC027353 Homo sapi
38	61.6	2.8	283769	2	AC094777	AC094777 Rattus no
c 40	60.6	2.8	250046	2	AC113592	AC113592 Homo sapi
41	60.4	2.8	300695	2	AC079431	AC079431 Mus muscu
42	60	2.8	60	6	CQ541600	CQ541600 Sequence
43	59.4	2.7	53552	2	AC100796	AC100796 Homo sapi
44	58.8	2.7	22466	2	AC136508	AC136508 Pan trogl
c 45	58.4	2.7	67246	2	AC105035	AC105035 Homo sapi

ALIGNMENTS

RESULT 1
AR071760
LOCUS AR071760 2180 bp DNA linear PAT 18-FEB-2000
DEFINITION Sequence 2 from patent US 5912142.
ACCESSION AR071760
VERSION AR071760.1 GI:7222648
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2180)
AUTHORS Kaufman,R.E. and Slentz-Kesler,K.
TITLE Gene product over expressed in cancer cells
JOURNAL Patent: US 5912142-A 2 15-JUN-1999;
FEATURES Location/Qualifiers
source
1..2180
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match	100.0%	Score	2180;	DB	6;	Length	2180;
Best Local Similarity	100.0%	Pred. No.	0;				
Matches	2180;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
Qy	1	ATTCTGCTTCTTTAGCGTGAACGCGGGTGGGTGCTCCCGTGAATAATAAATTCAC	60				
Db	1	ATTCTGCTTCTTTAGCGTGAACGCGGGTGGGTGCTCCCGTGAATAATAAATTCAC	60				
Qy	61	CGTACGGTCTGTGTGAACGCGGGTGGTTCGGAACCTTGAGGCTTCCCGTAAACCCAG	120				
Db	61	CGTACGGTCTGTGTGAACGCGGGTGGTTCGGAACCTTGAGGCTTCCCGTAAACCCAG	120				
Qy	121	CTCCTTCTCATCTGGGAGGTGGGTCCCGCGGGTGGGTCCCGGCTTCCCTGGCCCCCTC	180				
Db	121	CTCCTTCTCATCTGGGAGGTGGGTCCCGCGGGTGGGTCCCGGCTTCCCTGGCCCCCTC	180				
Qy	181	CCTCTCGTGTCTTTTCATTTCTTGGGGTCCGGGGCGGGAGAGCTGCATCCCCAGAGA	240				
Db	181	CCTCTCGTGTCTTTTCATTTCTTGGGGTCCGGGGCGGGAGAGCTGCATCCCCAGAGA	240				
Qy	241	CGCGTTCAGAGCGGACCGGGAGTGTTCAGAGCCAGTGACAGGACGAGGGGCCCA	300				
Db	241	CGCGTTCAGAGCGGACCGGGAGTGTTCAGAGCCAGTGACAGGACGAGGGGCCCA	300				

Db 241 GCGCGTCCAGGAGCGGACCCGGAGTGTTCAGAGCCAGTGACAAGGACGAGGGGCCA 300
Qy 301 AGTCCACAGCCATGAGACCTGCCCTCGCATTCCTGSCCAGCTTTCCAGGGCCCT 360
Db 301 AGTCCACAGCCATGAGACCTGCCCTCGCATTCCTGSCCAGCTTTCCAGGGCCCT 360
Qy 361 TGGGACCCCTCCCTGTTTGGCTGCTCCTCTGAGTGCTCAGAAATGAAGCTGGGACAGCC 420
Db 361 TGGGACCCCTCCCTGTTTGGCTGCTCCTCTGAGTGCTCAGAAATGAAGCTGGGACAGCC 420
Qy 421 CATCTGCAAGAGGGGTAGTCTCTGTGCTTGGGGCGAGACACCGCTCATGTCTGCAA 480
Db 421 CATCTGCAAGAGGGGTAGTCTCTGTGCTTGGGGCGAGACACCGCTCATGTCTGCAA 480
Qy 481 CATCTCAACGCTTCTCCCATGTCAACATCAAGCTGGGTGCCACGGGCGAGGAGCGC 540
Db 481 CATCTCAACGCTTCTCCCATGTCAACATCAAGCTGGGTGCCACGGGCGAGGAGCGC 540
Qy 541 CATCTTCAATGAGGTGGCTCCAGGCTACTTCTCCGGGAGCGCTGGCAGCTCCAGTTCA 600
Db 541 CATCTTCAATGAGGTGGCTCCAGGCTACTTCTCCGGGAGCGCTGGCAGCTCCAGTTCA 600
Qy 601 GGGAGGGTGGACAGCTGGTGATCAAGCGCGCGGACTCCCATGCTGGGCTGTACAT 660
Db 601 GGGAGGGTGGACAGCTGGTGATCAAGCGCGCGGACTCCCATGCTGGGCTGTACAT 660
Qy 661 GTGGCACTCTGGGACACAGAGAAATAACAGACAAAGTCAAGCTGGAGGTTTCAGTGCC 720
Db 661 GTGGCACTCTGGGACACAGAGAAATAACAGACAAAGTCAAGCTGGAGGTTTCAGTGCC 720
Qy 721 AGAACCCCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Db 721 AGAACCCCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Qy 781 CTTTCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
Db 781 CTTTCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
Qy 841 ACGCGGGAGAGAAAGTCTTCTCTCTAGAACCCAGATGAAGTGGCAGCCCTCAGAGC 900
Db 841 ACGCGGGAGAGAAAGTCTTCTCTCTAGAACCCAGATGAAGTGGCAGCCCTCAGAGC 900
Qy 901 GGGAGCCAGCAGGCGCTGAGCAGAGCTCGCTGAACTGTGGACCCCAAGCTCCAGGC 960
Db 901 GGGAGCCAGCAGGCGCTGAGCAGAGCTCGCTGAACTGTGGACCCCAAGCTCCAGGC 960
Qy 961 CACCCAAAGCGCTGGCACTGGTGTCAAAACCTCAACCTTGGAGCCCTGGAGCTGCT 1020
Db 961 CACCCAAAGCGCTGGCACTGGTGTCAAAACCTCAACCTTGGAGCCCTGGAGCTGCT 1020
Qy 1021 GTCCCCCAACCTTGTTCATATGCCGAGACCCATAGCCCTGCAAGGCAGAGAG 1080
Db 1021 GTCCCCCAACCTTGTTCATATGCCGAGACCCATAGCCCTGCAAGGCAGAGAG 1080
Qy 1081 GACACAGGAGACCAAGCCCTGAGTCCGACCTTGGGTGGCGGCGCTGGGCTCTCGTCC 1140
Db 1081 GACACAGGAGACCAAGCCCTGAGTCCGACCTTGGGTGGCGGCGCTGGGCTCTCGTCC 1140
Qy 1141 CACCCGAGGAGCAGACACCGGCTTGTGGAGGCTGGGCTCTGTGTCAACCACTCC 1200
Db 1141 CACCCGAGGAGCAGACACCGGCTTGTGGAGGCTGGGCTCTGTGTCAACCACTCC 1200
Qy 1201 TGGGTGCGTGACAGCCCTTCCCTCCACCCCGAGTCTTCCAGCTCTGTCTCAGT 1260
Db 1201 TGGGTGCGTGACAGCCCTTCCCTCCACCCCGAGTCTTCCAGCTCTGTCTCAGT 1260
Qy 1261 TTCCAAAATGGAACCACTCACTCTGAGACCCGATTTACAGGAGCATGCCCTCC 1320
Db 1261 TTCCAAAATGGAACCACTCACTCTGAGACCCGATTTACAGGAGCATGCCCTCC 1320
Qy 1321 CTCTGCCCTCATCAAAACCCAGACCCGAGTCTCCCTTCTGTCACCCCGAGGCTGGTCCGG 1380
Db 1321 CTCTGCCCTCATCAAAACCCAGACCCGAGTCTCCCTTCTGTCACCCCGAGGCTGGTCCGG 1380

Qy 1381 CCCCAGGTGGGGTCCGCTCTCTCACTCCAGGGCTCCGCGCCCAAGTGAAGGGGGCC 1440
Db 1381 CCCCAGGTGGGGTCCGCTCTCTCACTCCAGGGCTCCGCGCCCAAGTGAAGGGGGCC 1440
Qy 1441 CTGCGGAGCCTCAGACACACTGGATTGAGGCTGGGGGGGCTTGGACATACCTGTC 1500
Db 1441 CTGCGGAGCCTCAGACACACTGGATTGAGGCTGGGGGGGCTTGGACATACCTGTC 1500
Qy 1501 CTTGGCTATGACAGGCTTTGGGGCCCTTCCGGCGAGCCCGGGGGCCAGAGTAGGG 1560
Db 1501 CTTGGCTATGACAGGCTTTGGGGCCCTTCCGGCGAGCCCGGGGGCCAGAGTAGGG 1560
Qy 1561 TCTGGGGCTTAGAGGCTGGGATGGCTCTTGCGCCCAAGGGGGGCAAGCGCAGGCG 1620
Db 1561 TCTGGGGCTTAGAGGCTGGGATGGCTCTTGCGCCCAAGGGGGGCAAGCGCAGGCG 1620
Qy 1621 GGGCTGGAGGCGCGCGCGCTCGGGCTGGGGGTGAGGTGAGCGCTCCGCGGG 1680
Db 1621 GGGCTGGAGGCGCGCGCGCTCGGGCTGGGGGTGAGGTGAGCGCTCCGCGGG 1680
Qy 1681 CTGCTCCGCATCCCTCAGTCCCTCGGCCACCCGGGGGTGCTTCTGTCGCCACCGCA 1740
Db 1681 CTGCTCCGCATCCCTCAGTCCCTCGGCCACCCGGGGGTGCTTCTGTCGCCACCGCA 1740
Qy 1741 CTTGCCAGGCTCTTTGGACCCAGATCTGTTGATGCTTTTGTCTGTCACCTCGCGGG 1800
Db 1741 CTTGCCAGGCTCTTTGGACCCAGATCTGTTGATGCTTTTGTCTGTCACCTCGCGGG 1800
Qy 1801 GCGCTTTGATGCTTCTGATGAGTGGGTGGAATAATCACCGGGAATCCCTTTCAGTTC 1860
Db 1801 GCGCTTTGATGCTTCTGATGAGTGGGTGGAATAATCACCGGGAATCCCTTTCAGTTC 1860
Qy 1861 TTTGAAAAGTTCCATGACTCGAATATCTGAATGAAAGAAAACAAACCGACTCACAACC 1920
Db 1861 TTTGAAAAGTTCCATGACTCGAATATCTGAATGAAAGAAAACAAACCGACTCACAACC 1920
Qy 1921 TCCAAAGTCTCAAAATGCAATTTTAAATGGAATAACAAATCTGAAAGAAACGCTTT 1980
Db 1921 TCCAAAGTCTCAAAATGCAATTTTAAATGGAATAACAAATCTGAAAGAAACGCTTT 1980
Qy 1981 TAGTGGCTTTAAGCCCAAAACGCTCCCTAAGCGCTCTCGAGATGAAGACGGGGGGAGC 2040
Db 1981 TAGTGGCTTTAAGCCCAAAACGCTCCCTAAGCGCTCTCGAGATGAAGACGGGGGGAGC 2040
Qy 2041 CCGACCGAGGTGAGACCCCGCAGACCGGGGGGGCGGCTGACCGAGGCTTCGACAG 2100
Db 2041 CCGACCGAGGTGAGACCCCGCAGACCGGGGGGGCGGCTGACCGAGGCTTCGACAG 2100
Qy 2101 CCGGCGCCCTGAGGCTGGGCGGAGCGGCGGCGGCGGCTTCCAGAGGGGCGCTTGTCTCGG 2160
Db 2101 CCGGCGCCCTGAGGCTGGGCGGAGCGGCGGCGGCGGCTTCCAGAGGGGCGCTTGTCTCGG 2160
Qy 2161 TTAATAAAGTTCGCTCCG 2180
Db 2161 TTAATAAAGTTCGCTCCG 2180

RESULT 2
AR097657
LOCUS 2180 bp DNA linear PAT 14-FEB-2001
DEFINITION Sequence 2 from patent US 6072034.
ACCESSION AR097657
VERSION AR097657.1 GI:12806387
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2180)
AUTHORS Kaufman, R.E. and Slentz-Kesler, K.
TITLE Gene product over expressed in cancer cells
JOURNAL Patent: US 6072034-A 2 06-JUN-2000;
FEATURES Location/Qualifiers


```
Db      1621 GGGCTGGAGGGCGGCGGCGCTCGGGCTGGGGGTGAGTGGACGCTGCTCGGGG 1680
Qy      1681 CTGGTCGCGATCCCTCAGTCCCTCGGCCACCGGGGTCGCTCCCTCGTGGCCACCGCA 1740
Db      1681 CTGGTCGCGATCCCTCAGTCCCTCGGCCACCGGGGTCGCTCCCTCGTGGCCACCGCA 1740
Qy      1741 CTGGCGGAGCCTCTTTGGACCCAGATCTGTTTATGCTTTTGTCTGCTCACTCGCGCGG 1800
Db      1741 CTGGCGGAGCCTCTTTGGACCCAGATCTGTTTATGCTTTTGTCTGCTCACTCGCGCGG 1800
Qy      1801 GCGCTTTGATGCTTCATCTGATGGGTGGAAAAATCACCGGAATCCCGCTTCAGTTTC 1860
Db      1801 GCGCTTTGATGCTTCATCTGATGGGTGGAAAAATCACCGGAATCCCGCTTCAGTTTC 1860
Qy      1861 TTTGAAAAAGTTCCATGACTCGAATATCTGAATGAAGAAAAACAAACCGACTCACAAAC 1920
Db      1861 TTTGAAAAAGTTCCATGACTCGAATATCTGAATGAAGAAAAACAAACCGACTCACAAAC 1920
Qy      1921 TCCAGTAGCTCAAAATGCAATTTTAAATGGAAAAACAAAAATCTGAAAGAAAGCTCTT 1980
Db      1921 TCCAGTAGCTCAAAATGCAATTTTAAATGGAAAAACAAAAATCTGAAAGAAAGCTCTT 1980
Qy      1981 TAGTGGCTTTAAGCCCAAAAGCTCCCTAAGCGCTCTCGAGATGAAGACGGGGGAGC 2040
Db      1981 TAGTGGCTTTAAGCCCAAAAGCTCCCTAAGCGCTCTCGAGATGAAGACGGGGGAGC 2040
Qy      2041 CCCAGCAGGTGGAGACCCCGCAGGACGCGGGGCGCGGTGACCGAGGCTTCGCACAG 2100
Db      2041 CCCAGCAGGTGGAGACCCCGCAGGACGCGGGGCGCGGTGACCGAGGCTTCGCACAG 2100
Qy      2101 CCGGCGGCGCTGAGGGTCGGGCGGAGCCAGGGTCCAAAGGGGCGGTTTGTCTCGGG 2160
Db      2101 CCGGCGGCGCTGAGGGTCGGGCGGAGCCAGGGTCCAAAGGGGCGGTTTGTCTCGGG 2160
Qy      2161 TTTAAATAAGTTCGCTCCG 2180
Db      2161 TTTAAATAAGTTCGCTCCG 2180

RESULT 4
LOCUS      CQ776411          2000 bp      DNA      linear      PAT 11-MAR-2004
DEFINITION Sequence 97 from Patent EP1394274.
ACCESSION  CQ776411
VERSION    CQ776411.1  GI:45379801
KEYWORDS   .
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1
AUTHORS   Ohtani,N., Sugita,Y., Yamaya,M., Kubo,H., Nagai,H. and Izuwara,K.
TITLE     Methods of testing for bronchial asthma or chronic obstructive
            pulmonary disease
JOURNAL   Patent: EP 1394274-A 97 03-MAR-2004;
            Genex Research, Inc. (JP)
FEATURES   Location/Qualifiers
            source          1..2000
                        /organism="Homo sapiens"
                        /mol_type="unassigned DNA"
                        /db_xref="taxon:9606"
ORIGIN
Query Match      89.5%; Score 1951; DB 6; Length 2000;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1984; Conservative 0; Mismatches 0; Indels 3; Gaps 3;

Qy      196 ATTTTCCTGGGGTCGCGGGCGGAGAAAGCTGCATCCAGAGGAGCGGCTCCAGAGCG 255
Db      1 ATTTTCCTGGGGTCGCGGGCGGAGAAAGCTGCATCCAGAGGAGCGGCTCCAGAGCG 60
Qy      256 GACCCGGAGTGTTCAGAGCGAGTGAACAGGACAGGGGCCCAAGTCCCAACGACCAT 315
            |||||
```

```
Db      61  GACCGGGAGTGTTCCTCAAGAGCCAGTGAACAGGACGAGGGGCCCAAGTCCCACGACCAT 120
Qy      316  GCAGACCTGCCCCCTCGGCATTCCTTGGCCACGCTTTCCAGGGCCCTTGGAGCCCTCTCTGT 375
Db      121  GCAGACCTGCCCCCTCGGCATTCCTTGGCCACGCTTTCCAGGGCCCTTGGAGCCCTCTGT 180
Qy      376  TTTGGCTGCTCTCTTGTAGTGTTCAGAAAGAGGCTGGGACAGCCCCCATCTGCACAGAGGG 435
Db      181  TTTGGCTGCTCTCTTGTAGTGTTCAGAAAGAGGCTGGGACAGCCCCCATCTGCACAGAGGG 240
Qy      436  GGTAGTCTCTGTGTCTTGGGGCGGAGAACACCGTCTCATGTCTGCAACATCTCCAAAGCCCT 495
Db      241  GGTAGTCTCTGTGTCTTGGGGCGGAGAACACCGTCTCATGTCTGCAACATCTCCAAAGCCCT 300
Qy      496  CTCCCATGTCAACATCAAGCTCGCTGCCCAACCGGCGAGGAGCGCCCATCTTCAATGAGGT 555
Db      301  CTCCCATGTCAACATCAAGCTCGCTGCCCAACCGGCGAGGAGCGCCCATCTTCAATGAGGT 360
Qy      556  GGCTCAAGGCTACTTCTCCCGGACGGGTGGCAGTCTCAGGTTTCAGGGAGGGGTGGCACA 615
Db      361  GGCTCAAGGCTACTTCTCCCGGACGGGTGGCAGTCTCAGGTTTCAGGGAGGGGTGGCACA 420
Qy      616  GCTGTGATCAAGAGGCGCCCGGACCTCCCATGCTGGGCTGTACATGTGACATCTCGTGGG 675
Db      421  GCTGTGATCAAGAGGCGCCCGGACCTCCCATGCTGGGCTGTACATGTGACATCTCGTGGG 480
Qy      676  ACACAGAGAAATAACAGAACAGTCAACGCTGAGAGTTTCAGGTGCAGAAACCCAGTCCCG 735
Db      481  ACACAGAGAAATAACAGAACAGTCAACGCTGAGAGTTTCAGGTGCAGAAACCCAGTCCCG 540
Qy      736  CCTGACATGAGGTTCTGGCTGTGCCAGCGGTGTCACCTGTCTGCTTTCATCTCTTGGT 795
Db      541  CCTGACATGAGGTTCTGGCTGTGCCAGCGGTGTCACCTGTCTGCTTTCATCTCTTGGT 600
Qy      796  CGCTCTGCTCATGTTTCGCTGTGTAAGGTGCGCTGTTCCAGCAACCGCGGGAGAGAA 855
Db      601  CGCTCTGCTCATGTTTCGCTGTGTAAGGTGCGCTGTTCCAGCAACCGCGGGAGAGAA 660
Qy      856  GTTCTTCTCTTAGAACCCAGATGAAGGTCCAGCCCTCAGAGCGGAGCCAGCAGAGG 915
Db      661  GTTCTTCTCTTAGAACCCAGATGAAGGTCCAGCCCTCAGAGCGGAGCCAGCAGAGG 720
Qy      916  CTTGAGCAGAGCCTCGGTGAACTGTGACCCCAAGTCCCGAGCCCAACCCCAAGGCCGT 975
Db      721  CTTGAGCAGAGCCTCGGTGAACTGTGACCCCAAGTCCCGAGCCCAACCCCAAGGCCGT 780
Qy      976  GGCACCTGTTTCAAAACCTCACCACTGGAGCCCTGGAGCTGCTCTCCCGCCCAACCT 1035
Db      781  GGCACCTGTTTCAAAACCTCACCACTGGAGCCCTGGAGCTGCTGT - CCCCCCAACCT 839
Qy      1036  TGTTCCTATATGCCGAGACCCATAGCCGCTGCAAGGCGAGAGGAGACACAGGAGGCCA 1095
Db      840  TGTTCCTATATGCCGAGACCCATAGCCGCTGCAAGGCGAGAGGAGACACAGGAGGCCA 899
Qy      1096  GCGCTGAGTGCAGACCTTGGGTGGCGGGCTGGGTCTCTCTGTCACCCCGAGGGGACA 1155
Db      900  GCGCTGAGTGCAGACCTTGGGTGGCGGGCTGGGTCTCTCTGTCACCCCGAGGGGACA 959
Qy      1156  GACACGGCTTGTGGCAGGCTGGGCTCTGTGTTCACCACTCTGGGTGGGTGGCAGAC 1215
Db      960  GACACGGCTTGTGGCAGGCTGGGCTCTGTGTTCACCACTCTGGGTGGGTGGCAGAC 1019
Qy      1216  CTTTCCCTCCACCCCGCAGGTCTTCCAGCTCTGCTTCTCAGTTTCCAAATGGAACC 1275
Db      1020  CTTTCCCTCCACCCCGCAGGTCTTCCAGCTCTGCTTCTCAGTTTCCAAATGGAACC 1079
Qy      1276  ACCTCACTCTCGCAGACCCGACTTACAGGACGATGCCCCCTTCCCTCTGCGCTCATCAA 1335
Db      1080  ACCTCACTCTCGCAGACCCGACTTACAGGACGATGCCCCCTTCCCTCTGCGCTCATCAA 1139
Qy      1336  ACCCAGACCCGGACTCGCTTCTGTCACCCCGAGGCTGGTCCGGCCCCCAGGTGGGT 1395
Db      1140  ACCCAGACCCGGACTCGCTTCTGTCACCCCGAGGCTGGTCCGGCCCCCAGGTGGGT 1199
```

Wyeth (US); Burczynski, Michael E. (US); Twine, Natalie C. (US);
Dorner, Andrew J. (US); Trepicchio, William L. (US)
Location/Qualifiers
1. .2000
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

FEATURES

source

ORIGIN

Query Match 89.5%; Score 1951; DB 6; Length 2000;

Best Local Similarity 99.8%; Pred. No. 0; Indels 3; Gaps 3;
Matches 1984; Conservative 0; Mismatches 0

196 ATTTTCTCTGGGCTCCGGGCGCGGAGAGCTGCATCCAGAGGAGCGGTCCAGAGCG 255
Db 1 ATTTTCTCTGGGCTCCGGGCGCGGAGAGCTGCATCCAGAGGAGCGGTCCAGAGCG 60
256 GACCCGGAGTGTTCAGAGCCAGTGAAGAGCAGGGGCCCAAGTCCACCAAGCCAT 315
Db 61 GACCCGGAGTGTTCAGAGCCAGTGAAGAGCAGGGGCCCAAGTCCACCAAGCCAT 120
316 GCAGACCTGCCCTGGCATTCCTTGGCCACAGTTCCTCCAGGCCCTTGGGACCTCTCTGT 375
Db 121 GCAGACCTGCCCTGGCATTCCTTGGCCACAGTTCCTCCAGGCCCTTGGGACCTCTCTGT 180
376 TTTGGCTGCTCTTGAAGTGTCTCAGAAAGAGTGGGAGAGCCCATCTGCACAGAGG 435
Db 181 TTTGGCTGCTCTTGAAGTGTCTCAGAAAGAGTGGGAGAGCCCATCTGCACAGAGG 240
436 GGTAGTCTCTGTGTCTTGGGCGGAGAACACGTCATCTCTGCAACATCTTCAATAGGT 495
Db 241 GGTAGTCTCTGTGTCTTGGGCGGAGAACACGTCATCTCTGCAACATCTTCAATAGGT 300
496 CTCCCATGTCAACATCAAGTGTGCTGCGGCGGAGAGAGCCCATCTTCAATAGGT 555
Db 301 CTCCCATGTCAACATCAAGTGTGCTGCGGCGGAGAGAGCCCATCTTCAATAGGT 360
556 GGTCCAGGCTACTTCTCCGGGAGAGGCTGCGAGCTCCAGTTCAGGAGGCGTGACCA 615
Db 361 GGTCCAGGCTACTTCTCCGGGAGAGGCTGCGAGCTCCAGTTCAGGAGGCGTGACCA 420
616 GGTGTGTATCAAGGCGCGGAGCTCCCATCTCTGGGCTGTACATGTGGCACTCTGTGG 675
Db 421 GGTGTGTATCAAGGCGCGGAGCTCCCATCTCTGGGCTGTACATGTGGCACTCTGTGG 480
676 ACACAGAGAAATAACAGACAGTCAAGTTCAGGTTCAGGTTCAGGTTCAGGTTCAGGT 735
Db 481 ACACAGAGAAATAACAGACAGTCAAGTTCAGGTTCAGGTTCAGGTTCAGGTTCAGGT 540
736 CCTGACACTGGGTTCTGGCTGTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 795
Db 541 CCTGACACTGGGTTCTGGCTGTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
796 CGCTCTGGTCACTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 855
Db 601 CGCTCTGGTCACTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 660
856 GTTCTTCTCTAGAACCCAGTTCAGGTTCAGGTTCAGGTTCAGGTTCAGGTTCAGGTTCAGGT 915
Db 661 GTTCTTCTCTAGAACCCAGTTCAGGTTCAGGTTCAGGTTCAGGTTCAGGTTCAGGTTCAGGT 720
916 CCTGAGCAGAGCTCTGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 975
Db 721 CCTGAGCAGAGCTCTGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 780
976 GGCACCTGTGTTCAAAACCTTCAACCTTCAACCTTCAACCTTCAACCTTCAACCTTCAACCT 1035
Db 781 GGCACCTGTGTTCAAAACCTTCAACCTTCAACCTTCAACCTTCAACCTTCAACCTTCAACCT 839
1036 TGTTCCTATATCGGAGAGCCATAGCGCTGCAAGGAGGAGAGGAGGAGGAGGAGGAGGAGG 1095
Db 840 TGTTCCTATATCGGAGAGCCATAGCGCTGCAAGGAGGAGAGGAGGAGGAGGAGGAGGAGG 899

1396 CGCTCTCTCCACTCCAGGGCTCCGCCCAAGTAGGAGGGGCGCCCTCCGAGACCTCAG 1455
Db 1200 CGCTCTCTCTCCACTCCAGGGCTCCGCCCAAGTAGGAGGGGCGCCCTCCGAGACCTCAG 1259
1456 ACACACTGAGTTCAGAGGCTGGGGGGGCTTGGGCAATCTGTCTCTCTGTGCTATGAGCA 1515
Db 1260 ACACACTGAGTTCAGAGGCTGGGGGGGCTTGGGCAATCTGTCTCTCTGTGCTATGAGCA 1319
1516 GGTCTTGGGGGCTTCCGCGGAGGCTCCGCCCAAGTAGGAGGGGCTTGGGGCTTAGAG 1575
Db 1320 GGTCTTGGGGGCTTCCGCGGAGGCTCCGCCCAAGTAGGAGGGGCTTGGGGCTTAGAG 1379
1576 GCTGGGATGCTCTTGGGCTCCAGCGGAGGGGCTTGGGCAATCTGTCTCTCTGTGCTATGAG 1635
Db 1380 GCTGGGATGCTCTTGGGCTCCAGCGGAGGGGCTTGGGCAATCTGTCTCTCTGTGCTATGAG 1439
1636 GCGCGGCTCGGGCTTGGGGGCTTGGGCAATCTGTCTCTCTGTGCTATGAG 1695
Db 1440 GCGCGGCTCGGGCTTGGGGGCTTGGGCAATCTGTCTCTCTGTGCTATGAG 1499
1696 TCAGTCTCTCGGCTTGGGGGCTTGGGCAATCTGTCTCTCTGTGCTATGAG 1755
Db 1500 TCAGTCTCTCGGCTTGGGGGCTTGGGCAATCTGTCTCTCTGTGCTATGAG 1559
1756 TGGACCCAGATCTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGT 1815
Db 1560 TGGACCCAGATCTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGT 1619
1816 CATCTGTATGGGTGGGAAATCAACCGGAAATCCCGCTTCAAGTTCGAAAGTTCCA 1875
Db 1620 CATCTGTATGGGTGGGAAATCAACCGGAAATCCCGCTTCAAGTTCGAAAGTTCCA 1679
1876 TGACTCGAATATCTGAATGAAGAAACAAACCGACTCAAACTTCAAGTTCGAAAGTTCCA 1935
Db 1680 TGACTCGAATATCTGAATGAAGAAACAAACCGACTCAAACTTCAAGTTCGAAAGTTCCA 1739
1936 ATGCAATTTTAAATGGAACAAATCTGAAGAAAGTCTTGTAGTGGCTTTAAGCC 1995
Db 1740 ATGCAATTTTAAATGGAACAAATCTGAAGAAAGTCTTGTAGTGGCTTTAAGCC 1799
1996 CCAAACTCTTAAAGGCTCTCGAGTGAAGAGCGGGGGAG-CGCCAGCGAGTGA 2054
Db 1800 CCAAACTCTTAAAGGCTCTCGAGTGAAGAGCGGGGGAGCGCCAGCGAGTGA 1859
2055 GACCCCGAGGACCGGCGCGCGCTGACCGAGGCTTGCACAGCGCGCGCGCTGAG 2114
Db 1860 GACCCCGAGGACCGGCGCGCGCTGACCGAGGCTTGCACAGCGCGCGCGCTGAG 1919
2115 GGTGGGCTTCAAGAGGCGCGCTTGTGTCTCGGGTTAAATTAAGGTT 2173
Db 1920 GGTGGGCTTCAAGAGGCGCGCTTGTGTCTCGGGTTAAATTAAGGTT 1979
2174 CGGTCCG 2180
Db 1980 CGGTCCG 1986

RESULT 5
COB61464 2000 bp DNA linear PAT 10-SEP-2004
LOCUS
DEFINITION Sequence 97 from Patent WO2004072265.
ACCESSION COB61464
VERSION COB61464.1 GI:51982453
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 Burczynski, M., Twine, N., Dorner, A.J. and Trepicchio, W.L.
METHODS FOR MONITORING DRUG ACTIVITIES IN VIVO /i
Patent: WO 2004072265-A 97 26-AUG-2004;
JOURNAL

676 ACACAGAGAAATAACAGCAAGTCAAGCTGGAGGTTTTCAGGTGAGAAACCCAGTCCG 735
481 ACACAGAGAAATAACAGCAAGTCAAGCTGGAGGTTTTCAGGTGAGAAACCCAGTCCG 540
736 CCCTGACACTGGGTTCTGGCTGTGCGAGCGGTGTCTCTCTCTCTCTCTCTCTCTCTCT 795
541 CCCTGACACTGGGTTCTGGCTGTGCGAGCGGTGTCTCTCTCTCTCTCTCTCTCTCTCT 600
796 CGCTCTGGTTCATGTTTCGCTGTGATGAGTGGCTGTTTCCAGCAAGCCCGGAGAGAA 855
601 CGCTCTGGTTCATGTTTCGCTGTGATGAGTGGCTGTTTCCAGCAAGCCCGGAGAGAA 660
856 GTTCTTCTCTAGAACCCAGATGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 915
661 GTTCTTCTCTAGAACCCAGATGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 720
916 CTGAGCAGAGCTCGCTGAACTGTGACCCAGACTCCGAGCCCAACCCAGGCGCT 975
721 CTGAGCAGAGCTCGCTGAACTGTGACCCAGACTCCGAGCCCAACCCAGGCGCT 780
976 GGCACCTGGTTCACAAACCTCACCACTTGGAGCCCTGGAGTGTGTCTCTCTCTCTCTCT 1035
781 GGCACCTGGTTCACAAACCTCACCACTTGGAGCCCTGGAGTGTGTCTCTCTCTCTCTCT 839
1036 TGTTCATATGCGCAGACCCATAGCCGCTGCAAGGAGGAGGAGGAGGAGGAGGAGGAGG 1095
840 TGTTCATATGCGCAGACCCATAGCCGCTGCAAGGAGGAGGAGGAGGAGGAGGAGGAGG 899
1096 GGCCTGAGTCCGACCTTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG 1155
900 GGCCTGAGTCCGACCTTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG 959
1156 GACACCGGCTTGTGAGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 1215
960 GACACCGGCTTGTGAGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 1019
1216 CTTTCCCTCGACCCCGAGTCTTCAAGCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1275
1020 CTTTCCCTCGACCCCGAGTCTTCAAGCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1079
1276 ACCTCACTCCGACGACCCGACTTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1335
1080 ACCTCACTCCGACGACCCGACTTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1139
1336 ACCACAGACCGGACT 1395
1140 ACCACAGACCGGACT 1199
1396 CGGCTCTCTCACTCCGAGGCTCGGCGCCCAAGTGAAGGAGGAGGAGGAGGAGGAGGAGG 1455
1200 CGGCTCTCTCACTCCGAGGCTCGGCGCCCAAGTGAAGGAGGAGGAGGAGGAGGAGGAGG 1259
1456 ACACACTGAGTTACGGCTGGGGGGCTTTGGACATACCTGCTCTCTCTCTCTCTCTCTCTCT 1515
1260 ACACACTGAGTTACGGCTGGGGGGCTTTGGACATACCTGCTCTCTCTCTCTCTCTCTCTCT 1319
1516 GCGTTTGGGGGCTTCCGCGGACCCCGGGGCGAGGTAGGTCTGGGGGCTTGGAG 1575
1320 GCGTTTGGGGGCTTCCGCGGACCCCGGGGCGAGGTAGGTCTGGGGGCTTGGAG 1379
1576 GTTGGATGGCTCTTGGGCGCCACCGCAGGGGGCAAGCGCAGGCGGGCTGGAGGGCGG 1635
1380 GTTGGATGGCTCTTGGGCGCCACCGCAGGGGGCAAGCGCAGGCGGGCTGGAGGGCGG 1439
1636 GCGCGGCTCGGGGCTCAGGTGGAGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1695
1440 GCGCGGCTCGGGGCTCAGGTGGAGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1499
1696 TCAGTCCCTCGGCGACCCCGGGGCTCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1755
1500 TCAGTCCCTCGGCGACCCCGGGGCTCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1559
1756 TGGACCCAGATCTGTTTCATGCTTTTGTCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTT 1815

1560 TGGACCCAGATCTGTTTCATGCTTTTGTCTCTGCTCACTGCGCGGGGCCCTTTGATGCTT 1619
1816 CATCTGTATGGGGTGGAAAAATCAGCGGAATCCCTCTCAGTTCTTTTGAAGAGTCCA 1875
1620 CATCTGTATGGGGTGGAAAAATCAGCGGAATCCCTCTCAGTTCTTTTGAAGAGTCCA 1679
1876 TGAAGTGAATATCTGAAATGAAAGAAAAAAGAAACCGACTCAAAACCTCCAAAGTAGCTCAA 1935
1680 TGAAGTGAATATCTGAAATGAAAGAAAAAAGAAACCGACTCAAAACCTCCAAAGTAGCTCAA 1739
1936 ATGCAATTTTAAATATGAAAAAAGAAATCTGAAAGAAACGCTTTTAGTGGCTTTAAGCC 1995
1740 ATGCAATTTTAAATATGAAAAAAGAAATCTGAAAGAAACGCTTTTAGTGGCTTTAAGCC 1799
1996 CCAAAAGCTCCCTAAGCGCTCTCGAGATGAAAGCGGGGGGAG-CCCCAGGCCAGGTGA 2054
1800 CCAAAAGCTCCCTAAGCGCTCTCGAGATGAAAGCGGGGGGAGCCCCAGGCCAGGTGA 1859
2055 GACCCCGCAGGAGCGCGCGCGCTGACCGAGCTCGACAGCGCGCGCGCGCTGAG 2114
1860 GACCCCGCAGGAGCGCGCGCGCTGACCGAGCTCGACAGCGCGCGCGCGCTGAG 1919
2115 GGTGGGCGC-GAGCCAGGCTCCAGAGGGGCGGCTTGTCTCTCGGTTAAAAATAAGGTT 2173
1920 GGTGGGCGCAGGAGCGAGGCTCCAAAGAGGGGCGGCTTGTCTCTCGGTTAAAAATAAGGTT 1979
2174 CCGTCCG 2180
1980 CCGTCCG 1986

RESULT 7
AX333755 2000 bp DNA linear PAT 09-JAN-2002
LOCUS Sequence 4264 from Patent WO0194629.
DEFINITION AX333755
ACCESSION AX333755.1 GI:18124474
VERSION
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,
Horrigan, S., Soppet, D.R. and Weaver, Z.
TITLE Cancer gene determination and therapeutic screening using signature
gene sets
JOURNAL Patent: WO 0194629-A 4264 13-DEC-2001;
Avalon Pharmaceuticals (US)
FEATURES
source Location/Qualifiers
1..2000
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 89.5%; Score 1951; DB 6; Length 2000;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1984; Conservative 0; Mismatches 0; Indels 3; Gaps 3;
QY 196 ATTTTCTCTGGGCTCGGGCGCGGAGAGTGTCTCCAGAGGAGCGGTCCAGAGCG 255
DB 1 ATTTTCTCTGGGCTCGGGCGCGGAGAGTGTCTCCAGAGGAGCGGTCCAGAGCG 60
QY 256 GACCCGGGAGTGTTCAGAGCGGAGTGAAGACAGAGGGGCGGAGTCCACAGCCAT 315
DB 61 GACCCGGGAGTGTTCAGAGCGGAGTGAAGACAGAGGGGCGGAGTCCACAGCCAT 120
QY 316 GCAGACTCGGCTCGGCTTCTCTGCGCAAGTTCCTCCAGGCGCTTGGAGCCCTCTGTT 375
DB 121 GCAGACTCGGCTCGGCTTCTCTGCGCAAGTTCCTCCAGGCGCTTGGAGCCCTCTGTT 180

QY 376 TTTGGCTGCTCTGAGTGTCTAGAAATGAAGGCTGGGACAGCCCACTCTGCACAGAGG 435
DB 181 TTTGGCTGCTCTCTTTGAGTGTCTAGAAATGAAGGCTGGGACAGCCCACTCTGCACAGAGG 240
QY 436 GGTAGTCTCTGTGTCTTTGGGCGAGAAACACGCTCATGTCTTCAACATCTCCAAGCCCTT 495
DB 241 GGTAGTCTCTGTGTCTTTGGGCGAGAAACACGCTCATGTCTTCAACATCTCCAAGCCCTT 300
QY 496 CTCCCATGTCAACATCAAGCTCGTGTGCCACCGGCGAGGAGCGCCATCTTCAATGAGGT 555
DB 301 CTCCCATGTCAACATCAAGCTCGTGTGCCACCGGCGAGGAGCGCCATCTTCAATGAGGT 360
QY 556 GGTCTCAGGCTACTTCTCCGGGACGGCTGGGAGCTCCAGGTTTCAGGAGGCGTGGCACA 615
DB 361 GGTCTCAGGCTACTTCTCCGGGACGGCTGGGAGCTCCAGGTTTCAGGAGGCGTGGCACA 420
QY 616 GTGTGTGATCAAAAGGCGCGCGGACTCCCATGCTGGGCTGTACATGTGGACCTCGTGGG 675
DB 421 GTGTGTGATCAAAAGGCGCGCGGACTCCCATGCTGGGCTGTACATGTGGACCTCGTGGG 480
QY 676 ACACAGAGAAATAACAGAAAGTCAACGCTGGAGGTTTCAGGTGCAGAACCCAGTCCGC 735
DB 481 ACACAGAGAAATAACAGAAAGTCAACGCTGGAGGTTTCAGGTGCAGAACCCAGTCCGC 540
QY 736 CCTGACACTGGGTTCTGGCCTGTGCGCAGCGGTGTCTACTGTCTTCACTCTTGGT 795
DB 541 CCTGACACTGGGTTCTGGCCTGTGCGCAGCGGTGTCTACTGTCTTCACTCTTGGT 600
QY 796 CGCTCTGGTTCATGTTCCGCTGTACAGGTGCGCTGTTCCAGCAACGCCGGGAGAA 855
DB 601 CGCTCTGGTTCATGTTCCGCTGTACAGGTGCGCTGTTCCAGCAACGCCGGGAGAA 660
QY 856 GTTCTTCTCTTAGAACCCAGATGAAGTTCAGGCTCGAGCCCTTCAGAGCGGAGCCACAGGG 915
DB 661 GTTCTTCTCTTAGAACCCAGATGAAGTTCAGGCTCGAGCCCTTCAGAGCGGAGCCACAGGG 720
QY 916 CTTGACAGAGGCTCTCGTGAATCTGTGACCCAGACTTCGAGCCACCCCAAGCGCGT 975
DB 721 CTTGACAGAGGCTCTCGTGAATCTGTGACCCAGACTTCGAGCCACCCCAAGCGCGT 780
QY 976 GGCACCTGGTTCACAAACCTCACCACTTGGAGCCCTGGAGCTGTCTCCGCCCAACCT 1035
DB 781 GGCACCTGGTTCACAAACCTCACCACTTGGAGCCCTGGAGCTGTCTCCGCCCAACCT 839
QY 1036 TGTTCCTATATCCGAGACCCATAGCCGCTTCAGAGGAGAGGACACAGGAGGCA 1095
DB 840 TGTTCCTATATCCGAGACCCATAGCCGCTTCAGAGGAGAGGACACAGGAGGCA 899
QY 1096 GGCCTGAGTGGCAGACTTGGGTGGCGGGCTTGGGTCTCTGTCACACCCGAGGGCACA 1155
DB 900 GGCCTGAGTGGCAGACTTGGGTGGCGGGCTTGGGTCTCTGTCACACCCGAGGGCACA 959
QY 1156 GACACGCGCTGCTTGGCAGGCTGGGCTCTGTGTACCCACTCTGGGTGGTGCAGAC 1215
DB 960 GACACGCGCTGCTTGGCAGGCTGGGCTCTGTGTACCCACTCTGGGTGGTGCAGAC 1019
QY 1216 CTTTCCCTTCCACCCCGGAGTCTTCCAAAGCTCTGCTTCTCAGTTTCCAAATGGAAC 1275
DB 1020 CTTTCCCTTCCACCCCGGAGTCTTCCAAAGCTCTGCTTCTCAGTTTCCAAATGGAAC 1079
QY 1276 ACCTCACTCTCGAGACCCGACTTACAGAGAGCATGCCCTCTCTCTGCGCTCATCAA 1335
DB 1080 ACCTCACTCTCGAGACCCGACTTACAGAGAGCATGCCCTCTCTCTGCGCTCATCAA 1139
QY 1336 ACCACAGACCGGAGTCTCTTCTTCCACCCAGGCTGGTCCGCGCCAGGTGGGGT 1395
DB 1140 ACCACAGACCGGAGTCTCTTCTTCCACCCAGGCTGGTCCGCGCCAGGTGGGGT 1199
QY 1396 CGGCTCTCTTCACTCTCCAGGCTCTCGCGCCCAAGTGAAGGGGCGCTTGCAGGCTCAG 1455
DB 1200 CGGCTCTCTTCACTCTCCAGGCTCTCGCGCCCAAGTGAAGGGGCGCTTGCAGGCTCAG 1259
QY 1456 ACACACTGGAGTTTACAGGCTGGGGGGGCGCTTGGCACAATACCTGTCTCTTGGCTATGAGCA 1515

DB 1260 ACACACTGGAGTTTACAGGCTGGGGGGGCGCTTGGCACAATACCTGTCTCTTGGCTATGAGCA 1319
QY 1516 GGTCTTGGGGGCGCTTTCGCGGACGCCCGGGGCGGAGTAGGTCTGGGGGCTTAGAG 1575
DB 1320 GGTCTTGGGGGCGCTTTCGCGGACGCCCGGGGCGGAGTAGGTCTGGGGGCTTAGAG 1379
QY 1576 GCTGGATGGCTCTTGGGCCCAACCGGAGGGGCAAGCGAGCGCGGGCTGGAGGCGGC 1635
DB 1380 GCTGGATGGCTCTTGGGCCCAACCGGAGGGGCAAGCGAGCGCGGGCTGGAGGCGGC 1439
QY 1636 GCGCGGGCTCGGGCTGGGGGCTCAGGTGGAGCTGCTCCCTCGGGGCTGGTCCGGCATCC 1695
DB 1440 GCGCGGGCTCGGGCTGGGGGCTCAGGTGGAGCTGCTCCCTCGGGGCTGGTCCGGCATCC 1499
QY 1696 TCAGTCCCTCGGCGACCCCGGGGCTCGCTCCCTCGTGGCCACCGCACTTGCAGGCTCTT 1755
DB 1500 TCAGTCCCTCGGCGACCCCGGGGCTCGCTCCCTCGTGGCCACCGCACTTGCAGGCTCTT 1559
QY 1756 TGGACCCAGATCTGTTCATGTCTTTTGTCTTGTCTCACTCGGGGGGCGCTTTGATGTCTT 1815
DB 1560 TGGACCCAGATCTGTTCATGTCTTTTGTCTTGTCTCACTCGGGGGGCGCTTTGATGTCTT 1619
QY 1816 CATCTGTATGGGTGGAAAAATCACCGGGAATCCCGCTTCACTTCTTTGAAAAAGTTCCA 1875
DB 1620 CATCTGTATGGGTGGAAAAATCACCGGGAATCCCGCTTCACTTCTTTGAAAAAGTTCCA 1679
QY 1876 TGACTCGAATATCTGAAATGAAGAAAAAACAACCGACTCACAACTCTCAAGTAGTCCAA 1935
DB 1680 TGACTCGAATATCTGAAATGAAGAAAAAACAACCGACTCACAACTCTCAAGTAGTCCAA 1739
QY 1936 ATGCAATTTTAAATATGAAAAAACAATCTGAAAAAACAACGCTTTAGTGGCTTTAAGCC 1995
DB 1740 ATGCAATTTTAAATATGAAAAAACAATCTGAAAAAACAACGCTTTAGTGGCTTTAAGCC 1799
QY 1996 CCAAAACGCTCCCTAAGGCTCTCAGATGAAGACGGGGGGAG-CCCCAGCCAGGTGA 2054
DB 1800 CCAAAACGCTCCCTAAGGCTCTCAGATGAAGACGGGGGGAGCCCCCAGCCAGGTGA 1859
QY 2055 GACCCCGCAGACGCGCGCGCGCTGACGAGGCTCGCACAGCGCGCGCTGAG 2114
DB 1860 GACCCCGCAGACGCGCGCGCGCTGACGAGGCTCGCACAGCGCGCGCTGAG 1919
QY 2115 GGTCTGGGCTTCAAGCAGGCTTCAAGAGGGGCGGCTTTGTCTCGGTTAAATAAGGTT 2173
DB 1920 GGTCTGGGCTTCAAGCAGGCTTCAAGAGGGGCGGCTTTGTCTCGGTTAAATAAGGTT 1979
QY 2174 CGTCCG 2180
DB 1980 CGTCCG 1986

RESULT 8
HSU77643
LOCUS HSU77643 2000 bp mRNA linear PRI 10-MAR-1998
DEFINITION Homo sapiens K12 protein precursor mRNA, complete cds.
ACCESSION U77643
VERSION U77643.1 GI:2062390
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Slentz-Kesler, K.A., Hale, L.P. and Kaufman, R.E.
TITLE Identification and characterization of K12 (SCTM1), a novel human gene that encodes a Golgi-associated protein with transmembrane and secreted isoforms
JOURNAL Genomics 47 (3), 327-340 (1998)
MEDLINE 98149980
PubMed 9480746
REFERENCE
AUTHORS Slentz-Kesler, K.A. and Kaufman, R.E.

Sun Mar 20 14:17:43 2005

TITLE Direct Submission
Submitted (07-NOV-1996) Biochemistry, Duke University, Box 3250
JOURNAL DUMC, Durham, NC 27710, USA
FEATURES
source
1..2000
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="17"
/map="17q25.2-25.3"
/cell_type="K562 erythroleukemic"
119..865
/note="type Ia transmembrane protein"
CDS
/codon_start=1
/product="K12 protein precursor"
/protein_id="AAC52044.1"
/db_xref="GI:2062391"
/translation="MQTCFLAPFPHVSHQALGTLFLAASLSAQNEGWDSPICTEGVSVS
VSWGENTVMSCNISAFNSHVNPKLRAHGOESAIENEVAPGFPSRDGWLQVQGGVAOL
VIKGARDSHAGIYMHVLVGHQRNRRQVTLVSGAEPQAPDPTGFVPAVIAVFILL
VALVMPAWYRCRCSQORREKFFLEPQMKVAALRAQAQGLSRASABELWTPDSEPTP
RPLALVFKPSPLGALELLSPQFLFPVADP"
119..202
sig_peptide
mat_peptide
203..862
/product="K12 protein"
misc_feature
551..623
/note="encodes transmembrane domain"
misc_feature
752..862
/note="encodes proline-rich region"
ORIGIN
Query Match 89.5%; Score 1951; DB 9; Length 2000;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1984; Conservative 0; Mismatches 0; Indels 3; Gaps 3;
QY 196 ATTTTCCTGGGCTCGGGCGGGAGAGCTGCATCCAGAGAGCGGCTCCAGAGCG 255
DB 1 ATTTTCCTGGGCTCGGGCGGGAGAGCTGCATCCAGAGAGCGGCTCCAGAGCG 60
QY 256 GACCGGGAGTGTTCAGAGCCAGTGAAGAGCAGAGGCGCCCAAGTCCACAGCCAT 315
DB 61 GACCGGGAGTGTTCAGAGCCAGTGAAGAGCAGAGGCGCCCAAGTCCACAGCCAT 120
QY 316 GCAGACCTGCCCTGGCATTCCTGGCCAGCTTCCAGGCGCTTGGGACCTCCTGTT 375
DB 121 GCAGACCTGCCCTGGCATTCCTGGCCAGCTTCCAGGCGCTTGGGACCTCCTGTT 180
QY 376 TTTGGCTCCCTTCAGTGTCTCAGAATGAAGCTGGGACAGCGCCCATCTGCAAGAGG 435
DB 181 TTTGGCTCCCTTCAGTGTCTCAGAATGAAGCTGGGACAGCGCCCATCTGCAAGAGG 240
QY 436 GGTAGTCTGTGTCTGGGCGGAGAACACCGTGTATGCTTGCACATCTCCAGCGCTT 495
DB 241 GGTAGTCTGTGTCTGGGCGGAGAACACCGTGTATGCTTGCACATCTCCAGCGCTT 300
QY 496 CTCCTCATGTCAACATCAAGCTCGTGCACCGGCGAGAGAGCGCCCATCTTCAATGAGGT 555
DB 301 CTCCTCATGTCAACATCAAGCTCGTGCACCGGCGAGAGAGCGCCCATCTTCAATGAGGT 360
QY 556 GGCTCCAGGCTACTTCTCCCGGAGCGGTGGAGCTCCAGGTCAGGGAGCGTGGACACA 615
DB 361 GGCTCCAGGCTACTTCTCCCGGAGCGGTGGAGCTCCAGGTCAGGGAGCGTGGACACA 420
QY 616 GTGTGTGATCAAGAGCGCGGAGCTCCCATGCTGGGCTGTACATGTGGCACCTCGTGGG 675
DB 421 GTGTGTGATCAAGAGCGCGGAGCTCCCATGCTGGGCTGTACATGTGGCACCTCGTGGG 480
QY 676 ACACAGAGAAATAACAGACAAAGTCAAGTGGAGGTTTCAGTGCAGAAACCCAGTCCGC 735
DB 481 ACACAGAGAAATAACAGACAAAGTCAAGTGGAGGTTTCAGTGCAGAAACCCAGTCCGC 540
QY 736 CCTGTGACTGGGTTCTGGCCCTGTGCAGCGGTGTCTCACTGCTGTCTTCATCCTCTTGGT 795

Db
541 CCCTGACACTGGGTTCTGGCCTGTGCCAGCGGTGGTCACTGCTGTCTTCACTCCTCTTGGT 600
QY
796 CGCTCTGGTCACTGTTTCGCTGTGTACAGGTGCGGCTGTTCCTCCAGCAACGCCGCGGAGAA 855
Db
601 CGCTCTGGTCACTGTTTCGCTGTGTACAGGTGCGGCTGTTCCTCCAGCAACGCCGCGGAGAA 660
QY
856 GTTCTTCTCTCTAGAAACCCAGATGAAGTTCGAGCCCTCAGAGCGGGAGCCAGCAGGG 915
Db
661 GTTCTTCTCTCTAGAAACCCAGATGAAGTTCGAGCCCTCAGAGCGGGAGCCAGCAGGG 720
QY
916 CTTGAGCAGAGCCTTCGCTGAACTGTGACCCCAAGACTCCGAGCCCAACCCCAAGGCCGCT 975
Db
721 CTTGAGCAGAGCCTTCGCTGAACTGTGACCCCAAGACTCCGAGCCCAACCCCAAGGCCGCT 780
QY
976 GGCACCTGGTGTCAAAACCCCTCAACACTTGGAGCCCTGGAGCTGTCTGCTCCCCCAACCCCT 1035
Db
781 GGCACCTGGTGTCAAAACCCCTCAACACTTGGAGCCCTGGAGCTGTCTGCTCCCCCAACCCCT 839
QY
1036 TGTTCATATGCGCAGAGCCCATAGCCGCTCTCAAGGCGAGAGACACACAGAGAGCCA 1095
Db
840 TGTTCATATGCGCAGAGCCCATAGCCGCTCTCAAGGCGAGAGACACACAGAGAGCCA 899
QY
1096 GCCTGTAGTCCGACCTTGGTGGCGGGCTCGGTCTCTCTGTCCTCAACCCCGAGAGGACACA 1155
Db
900 GCCTGTAGTCCGACCTTGGTGGCGGGCTCGGTCTCTCTGTCCTCAACCCCGAGAGGACACA 959
QY
1156 GACACCGGCTTGTTCGAGGCTGGGCTCTGTGTCAACCACTCTGGGTCGTGTGACGAC 1215
Db
960 GACACCGGCTTGTTCGAGGCTGGGCTCTGTGTCAACCACTCTCTGGGTCGTGTGACGAC 1019
QY
1216 CTTTCCTCTCACCCCGCAGCTTTCAGCTCTGCTTCTCAGTTTCCAAATATGGAAC 1275
Db
1020 CTTTCCTCTCACCCCGCAGCTTTCAGCTCTGCTTCTCAGTTTCCAAATATGGAAC 1079
QY
1276 ACCTCACCTCCGAGCAGCCGACTTACAGAGCCATGCCCCCTCTCTGCTCTCATCAA 1335
Db
1080 ACCTCACCTCCGAGCAGCCGACTTACAGAGCCATGCCCCCTCTCTGCTCTCATCAA 1139
QY
1336 ACCACAGACCCGAGCTCTCTTTCGCAACCCAGGCTGTCCGCGCCAGGTCGSGGT 1395
Db
1140 ACCACAGACCCGAGCTCTCTTTCGCAACCCAGGCTGTCCGCGCCAGGTCGSGGT 1199
QY
1396 CGGCTCTCTCACCTCCAGAGGCTCGCGCCCAAGTAGAGGGGCGCCCTGCGGAGCCTCAG 1455
Db
1200 CGGCTCTCTCACCTCCAGAGGCTCGCGCCCAAGTAGAGGGGCGCCCTGCGGAGCCTCAG 1259
QY
1456 ACACACTGGAGTTCAAGGCTGGGGGGCTTGGGACATACCTGTCTCTGCTATGAGCA 1515
Db
1260 ACACACTGGAGTTCAAGGCTGGGGGGCTTGGGACATACCTGTCTCTGCTATGAGCA 1319
QY
1516 GGCTTTGGGGGCTTTCGCGGCGAGCCCGGGGGCGAGGTAGGCTCTGGGGGCTTAGAG 1575
Db
1320 GGCTTTGGGGGCTTTCGCGGCGAGCCCGGGGGCGAGGTAGGCTCTGGGGGCTTAGAG 1379
QY
1576 GCTGGATGGCTCTCTGCGCCCAACCCAGGGGGCAAGCGCAGGCGCGGCTGGAGGCGGC 1635
Db
1380 GCTGGATGGCTCTCTGCGCCCAACCCAGGGGGCAAGCGCAGGCGCGGCTGGAGGCGGC 1439
QY
1636 GCGCGGCGCTCGGGGCTGAGGTGGAGCGCTTCCCGGGGCTGTGCGCGCATCC 1695
Db
1440 GCGCGGCGCTCGGGGCTGAGGTGGAGCGCTTCCCGGGGCTGTGCGCGCATCC 1499
QY
1696 TCACTCTCGGCGCCCGGGGGTGGCTCTCTGTCGCCCAACCCGACCTGCGGAGCCTCTT 1755
Db
1500 TCACTCTCGGCGCCCGGGGGTGGCTCTCTGTCGCCCAACCCGACCTGCGGAGCCTCTT 1559
QY
1756 TGGACCCAGATCTGTTTCATGCTTTTGTCTGTCTGCTGCTGCGGGGGGCTTTGATGCTT 1815
Db
1560 TGGACCCAGATCTGTTTCATGCTTTTGTCTGCTGCTGCTGCGGGGGGCTTTGATGCTT 1619
QY
1816 CATCTGTATGGGGTGGAAAAATCACCGGGAATCCCCCTTTCAGTTCTTTTGAAGAGTTCCA 1875
Db
1620 CATCTGTATGGGGTGGAAAAATCACCGGGAATCCCCCTTTCAGTTCTTTTGAAGAGTTCCA 1679

1876 TGACTCGAATATCTGAATGAAGAAACAAACCGACTCACAACCTTCAAGTAGCTCCAA 1935
 1680 TGACTCGAATATCTGAATGAAGAAACAAACCGACTCACAACCTTCAAGTAGCTCCAA 1739
 1936 ATGCAATTTTAAATGGAACCAAAATCTGAAGAAACGCTTTAGTGGCTTTAAGCC 1995
 1740 ATGCAATTTTAAATGGAACCAAAATCTGAAGAAACGCTTTAGTGGCTTTAAGCC 1799
 1996 CCAAAACGTCCTTAAGGCGCTCTCGAGATGAAGACGGGGGGGAG-CCCCAGGACAGTGGA 2054
 1800 CCAAAACGTCCTTAAGGCGCTCTCGAGATGAAGACGGGGGGGAGCCCCAGGACAGTGGA 1859
 2055 GACCCCGCAGACGCGGCGCGCGGTGACCGAGCGCTTCGCACAGCGGCGCCCTGAG 2114
 1860 GACCCCGCAGACGCGGCGCGCGGTGACCGAGCGCTTCGCACAGCGGCGCCCTGAG 1919
 2115 GGTGCGGCGC-GAGCCAGAGGTCAAGAGGGGCGCGTTCGTCTCGGTAAATAGGTT 2173
 1920 GGTGCGGCGCGGAGCCAGGCGTCAAGAGGGGCGCGTTTGTCTCGGTTAAATAGGTT 1979
 2174 CCGTCCG 2180
 1980 CCGTCCG 1986

BC017716 1761 bp mRNA linear PRI 29-JUN-2004
 Homo sapiens secreted and transmembrane 1, mRNA (cdna clone
 MGC:21265 IMAGE:4413387), complete cds.

BC017716
 MGC:
 BC017716.1 GI:17389345
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1761)
 Strausberg,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
 Altschul,S.P., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
 Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
 Diatchenko,L., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
 Stapleton,M., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
 Scheetz,T.E., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
 Abramson,R.D., Mullany,S.J., Bosak,S.A., McEwan,P.J.,
 McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
 Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
 Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
 Fahey,J., Helton,E., Kettman,M., Madan,A., Young,A.C., Rodrigues,S.,
 Bouffard,G.G., Blakeley,R.W., Touchman,J.W., Green,E.D.,
 Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
 Butlerfield,Y.S., Krzywinski,M.I., Skalek,U., Smailus,D.E.,
 Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 12477932
 2 (bases 1 to 1761)
 Strausberg,R.
 Direct Submission
 Submitted (03-DEC-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: ATCC
 cdna Library Preparation: Life Technologies, Inc.

cdna Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome
 Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www-shgc.stanford.edu>
 Contact: (Dickson, Mark) mcdopaxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
 R. M.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAK Plate: 27 Row: h Column: 10
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 4506868.
 Location/Qualifiers
 1. 1761
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:21265 IMAGE:4413387"
 /tissue_type="Liver, adenocarcinoma"
 /clone_lib="NIH MGC_90"
 /lab_host="DH10B"
 /note="vector: pCMV-SPORT6"
 1. 1761
 /gene="SECTM1"
 /note="synonym: K12"
 /db_xref="LocusID:6398"
 /db_xref="MIM:602602"
 93..839
 /gene="SECTM1"
 /codon_start=1
 /product="secreted and transmembrane 1 precursor"
 /protein_id="AAH17716.1"
 /db_xref="GI:17389346"
 /db_xref="LocusID:6398"
 /db_xref="MIM:602602"
 /translation="MQTCPLAFPGHVSQLGTLFLAASLSAQNWGSDPCTEGVVS
 VSWGENTVMSCNISNASHVNIKLRAHQESAIFNEVAPGYFSRDGWLQVGGVAQL
 VIKGARDSHAGLYMHWHLVGHQNNRQVTLVSGAEPQSDPTGFWPVPVAVTVALL
 VALVWFAYRCRCQORREKFFLELPOMKFAALRAQAQGLSRASAEIWLTPDSPTP
 RPLALVFKPSPLGALLELLSPQLFPYADP"

ORIGIN
 Query Match 77.1%; Score 1681.2; DB 9; Length 1761;
 Best Local Similarity 99.2%; Pred. No. 0;
 Matches 1742; Conservative 0; Mismatches 8; Indels 6; Gaps 5;
 QY 222 GAAGCTGATCCAGAGGAGCGCTCCAGGAGCGGACCCGGAGTGTTCAGAGCCAGT 281
 DB 1 GAAGCTGATCCAGAGGAGCGCTCCAGGAGCGGACCCGGAGTGTTCAGAGCCAGT 60
 QY 282 GACAGGACAGGGGCCCAAGTCCCAACAGCATGTCAGACCTGCCCCCTGGCATTCCTCG 341
 DB 61 GACAGGACAGGGGCCCAAGTCCCAACAGCATGTCAGACCTGCCCCCTGGCATTCCTCG 120
 QY 342 GCCAGTTTCCAGGCCCTTGGAGCCCTCTCTGTTTTGGCTGCTCTTGAAGTGTCTCAGA 401
 DB 121 GCCAGTTTCCAGGCCCTTGGAGCCCTCTCTGTTTTGGCTGCTCTTGAAGTGTCTCAGA 180
 QY 402 ATGAAGGCTGGGACAGCCCCATCTGCACAGAGGGGTAGTCTCTGTCTTGGGGCGAGA 461
 DB 181 ATGAAGGCTGGGACAGCCCCATCTGCACAGAGGGGTAGTCTCTGTCTTGGGGCGAGA 240
 QY 462 ACACGCTCATGCTCTGCAACATCTCCAACGCTCTCTCCATGTCAACATCAAGCTCGCGT 521
 DB 241 ACACGCTCATGCTCTGCAACATCTCCAACGCTCTCTCCATGTCAACATCAAGCTCGCGT 300
 QY 522 CCCACGGGACAGAGCGCCATCTTCAATAGGTTGGCTCCAGGCTACTTCTCCCGGAGG 581
 DB 301 CCCACGGGACAGAGCGCCATCTTCAATAGGTTGGCTCCAGGCTACTTCTCCCGGAGG 360
 QY 582 GCTGGAGCTCCAGTTCAAGGAGCGGTGGACAGCTGGTGTATCAAGGCGCGCCGGAGT 641

361 GCTGGCAGCTCCAGTTTCAGGAGGCGGTGGCA CAGCTGGTGTATCAAAAGGCGCCCGGACT 420
Db
642 CCCATGCTGGGCTGTATCATGTGGACCTCGTGGGACACACAGAGAAATAACAGACAAGTCA 701
Qy
421 CCCATGCTGGGCTGTATCATGTGGACCTCGTGGGACACACAGAGAAATAACAGACAAGTCA 480
Db
702 CGCTGAGGTTTCAGTGCAGAAACCCAGTCCGCCCTTGACACTGGGTTCTGGGCTGTGC 761
Qy
481 CGCTGAGGTTTCAGTGCAGAAACCCAGTCCGCCCTTGACACTGGGTTCTGGGCTGTGC 540
Db
762 CAGCGTGTGTATCATGCTGTTTCATCTCTTGGTGCCTCTGGTCAATGTTCCGCTGGTACA 821
Qy
541 CAGCGTGTGTATCATGCTGTTTCATCTCTTGGTGCCTCTGGTCAATGTTCCGCTGGTACA 600
Db
822 GGTGGCGCTGTTCCAGCAACCCGCGGAGAAAGTTCCTCTCTAGAACCCAGATGA 881
Qy
601 GGTGGCGCTGTTCCAGCAACCCGCGGAGAAAGTTCCTCTCTAGAACCCAGATGA 660
Db
882 AGTTCGAGCCCTCAGAGCGGAGCCAGCAGAGGCTGTAGCAGAGCTCCGCTGAACGTGT 941
Qy
661 AGTTCGAGCCCTCAGAGCGGAGCCAGCAGAGGCTGTAGCAGAGCTCCGCTGAACGTGT 720
Db
942 GGAACCCAGACTCCGAGGCGCCACCCCAAGCGCTGGCACTGGTGTCAAAACCTCACCAC 1001
Qy
721 GGAACCCAGACTCCGAGGCGCCACCCCAAGCGCTGGCACTGGTGTCAAAACCTCACCAC 780
Db
1002 TTGGAGCCCTGGAGCTGTCTCCGCCCAACCCCTGTTTCCATATGTCGCGAGACCATAG 1061
Qy
781 TTGGAGCCCTGGAGCTGTCTCCGCCCAACCCCTGTTTCCATATGTCGCGAGACCATAG 839
Db
1062 CCGCTTCAGAGCAGAGAGACACAGAGGAGCCCTGAGTGGGCTGGTGGG 1121
Qy
840 CCGCTTCAGAGCAGAGAGACACAGAGGAGCCCTGAGTGGGCTGGTGGG 899
Db
1122 GGGCTCGGCTCTCTGCTCCACCGAGGCGCAGACACCGGCTGCTTGGCAGGCTGG 1181
Qy
900 GGGCTCGGCTCTCTGCTCCACCGAGGCGCAGACACCGGCTGCTTGGCAGGCTGG 959
Db
1182 CCTGTGTATCACCACCTCTGGGTGTGAGACCCCTTCCCTCCACCCCGAGGCTTC 1241
Qy
960 CCTGTGTATCACCACCTCTGGGTGTGAGACCCCTTCCCTCCACCCCGAGGCTTC 1019
Db
1242 CAGCTCTGCTCTCTCAGTTTCCAAATGGAACCACTCCTCCGCGAGACCCGACTTA 1301
Qy
1020 CAGCTCTGCTCTCTCAGTTTCCAAATGGAACCACTCCTCCGCGAGACCCGACTTA 1079
Db
1302 CCAGGACGATGCGCCCTCTGCTCTCATCAAAACCGACAGCCCGACTTCTGCTG 1361
Qy
1080 CCAGGACGATGCGCCCTCTGCTCTCATCAAAACCGACAGCCCGACTTCTGCTG 1139
Db
1362 CCAGCCAGGCTGTGTCGGGCGCCAGGTGTGGGTCGCTCTCTCCTCCTCCAGGCTCCG 1421
Qy
1140 CCAGCCAGGCTGTGTCGGGCGCCAGGTGTGGGTCGCTCTCTCCTCCTCCAGGCTCCG 1199
Db
1422 CCAGCCAGGCTGTGTCGGGCGCCAGGTGTGGGTCGCTCTCTCCTCCTCCAGGCTCCG 1479
Qy
1200 CCAGCCAGGCTGTGTCGGGCGCCAGGTGTGGGTCGCTCTCTCCTCCTCCAGGCTCCG 1259
Db
1480 GGGCTTGGCAATACCTGCTGCTGCTGATGAGCAGGCTTGGGCGCCCTTCCGCGGCA 1539
Qy
1260 GGGCTTGGCAATACCTGCTGCTGCTGATGAGCAGGCTTGGGCGCCCTTCCGCGGCA 1319
Db
1540 GGGCGGGGCGAGGTAGGCTGCTGGGCGCTTGGAGGCTGGGATGGCTCTGGGCGCCAC 1599
Qy
1320 GGGCGGGGCGAGGTAGGCTGCTGGGCGCTTGGAGGCTGGGATGGCTCTGGGCGCCAC 1378
Db
1600 GCGAGGGGCAAGCGAGGCGGGCTGGAGGCGGGGCGGCGGCTCGGGCTGGGGGCTC 1659
Qy
1379 GCGAGGGGCGAGCGAGGCGGGCTGGAGGCGGGGCGGCGGCTCGGGCTGGGGGCTC 1437
Db
1660 AGGTGGAGCGCTGCTCGGGGCTGGTGGCGGCTGCTCAGTCCCTCGGCGCAACCGGGG 1718
Qy
1438 AGGTGGAGCGCTGCTCGGGGCTGGTGGCGGCTGCTCAGTCCCTCGGCGCAACCGGGG 1497
Db

1719 TCGCTCCCTCGTCCGACCGACCTCGGAGCTCTTTGGACCCAGATCTGTATGCTT 1778
Qy
1498 TCGCTCCCTCGTCCGACCGACCTCGGAGCTCTTTGGACCCAGATCTGTATGCTT 1557
Db
1779 TTGTCTTCCTCACTGCGGCGGCGCTTTGATGTCTTCATCTGTATGGGTGGAATATC 1838
Qy
1558 TTGTCTTCCTCACTGCGGCGGCGCTTTGATGTCTTCATCTGTATGGGTGGAATATC 1617
Db
1839 ACCGGGAATCCCGCTTCAGTTCTTTGAAAAGTCCATGACTCGAATATCTGAAATGAAG 1898
Qy
1618 ACCGGGAATCCCGCTTCAGTTCTTTGAAAAGTTCATGACTCGAATATCTGAAATGAAG 1677
Db
1899 AAAACAAACCGACTCACAACCTCCAAAGTAGCTCCAAATGCAATTTTAAATGGAATAC 1958
Qy
1678 AAAACAAACCGACTCACAACCTCCAAAGTAGCTCCAAATGCAATTTTAAATGGAATAC 1737
Db
1959 AAAAATCTGAAAGAAA 1974
Qy
1738 AAAAATCTGAAAGAAA 1753
Db

RESULT 10
AC132938 154252 bp DNA linear PRI 20-FEB-2003
LOCUS Homo sapiens chromosome 17, clone RP13-20L14, complete sequence.
DEFINITION AC132938
ACCESSION AC132938.9 GI:28436295
VERSION HTG.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Bukaryota; Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 154252)
AUTHORS Birren, B., Nusbaum, C. and Lander, E.
TITLE Homo sapiens chromosome 17, clone RP13-20L14
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 154252)
AUTHORS Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,
Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B.,
Camara, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A.,
Cook, A., Cooke, P., DeArrellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J.,
Gardyna, S., Gord, S., Graham, L., Grand-pierre, N., Hagos, B.,
Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kanat, A.,
Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K.,
Liu, G., MacLean, C., Macdonald, P., Major, J., Matthews, C.,
McCarthy, M., Meldrim, J., Meneus, L., Minova, T., Miura, C., Norman, C.H.,
Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Peterson, K.,
O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Rise, C., Rogov, P.,
Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Seaman, S., Severy, P.,
Roman, J., Roy, A., Schauer, S., Schupback, R., Stojanovic, N., Talamas, J.,
Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J.,
Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J.,
Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (04-SEP-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 154252)
Birken, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N.,
Anderson, S., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T.,
Boguslavsky, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y.,
Collymore, A., Cooke, P., Cooke, P., Corum, B., DeArrellano, K.,
Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S.,
Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S.,
Graham, L., Grand-pierre, N., Hafez, N., Hagopian, D., Hagos, B.,
Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R.,
Lindblad-Toh, K., Liu, G., Lui, A., Mabbitt, R., MacLean, C.,
McCarthy, M., Major, J., Manning, J., Matthews, C., McCarthy, M.,
Meldrim, J., Meneus, L., Minova, T., Miura, C., Norman, C.H.,
Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,
O'Neil, D., Oliver, J., Rise, C., Rogov, P.,
Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Seaman, S., Severy, P.,
Roman, J., Roy, A., Schauer, S., Schupback, R., Stojanovic, N., Talamas, J.,
Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J.,
Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J.,
Zembek, L., Zimmer, A. and Zody, M.

O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramsamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schuback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (31-JAN-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 154252)

Anderson, S., Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N., Boguslavsky, L., Boukagalter, B., Camarata, J., Chang, J., Choepel, X., Collumore, A., Cook, A., Cooke, P., Corum, B., DeArellano, K., Diaz, J. S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Lui, A., Mabbitt, R., MacLean, C., MacDonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramsamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schuback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (20-FEB-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Feb 20, 2003 this sequence version replaced gi:28173196.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

TITLE JOURNAL COMMENT

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu/>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L27988
Center clone name: 20_L_14

Only the first 154.3 kilobases of this clone are being submitted.
The remainder overlaps accession number AC124287 [WICGR project L27363].

FEATURES source

Location/Qualifiers
1..154252
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosomes="17"
/map="17"
/clone="RP13-20L14"
/clone_lib="RP13-13 Human Female BAC"
1..32
repeat_region
167..473
/rpt_family="Alu"
repeat_region
490..727
/rpt_family="AluSx"
repeat_region
1062..1336
/rpt_family="L1ME"
repeat_region
1611..1974
/rpt_family="AluJb"
repeat_region
3544..3664
/rpt_family="MSTA"
repeat_region
complement(1392..3954)
/rpt_family="L2"

repeat_region
complement(4782..4945)
/rpt_family="MIR"
5143..5294
/rpt_family="MIR"
complement(5298..5589)
/rpt_family="AluSx"
5696..6016
/rpt_family="AluJb"
complement(6156..6436)
/rpt_family="AluY"
complement(6445..6733)
/rpt_family="AluSx"
complement(7025..7336)
/rpt_family="AluSx"
9135..9432
/rpt_family="MER72"
9440..9721
/rpt_family="MER72"
complement(9727..9881)
/rpt_family="L1R8A"
complement(9882..10192)
/rpt_family="AluSg"
complement(10193..10749)
/rpt_family="L1R8A"
complement(10750..11048)
/rpt_family="AluJb"
complement(11049..11084)
/rpt_family="L1R8A"
11108..11399
/rpt_family="AluSx"
11598..11912
/rpt_family="AluSg"
11946..12247
/rpt_family="AluSx"
complement(12259..12412)
/rpt_family="L1MCS"
12462..13718
/rpt_family="SVA"
13781..14081
/rpt_family="AluSx"
complement(14198..14852)
/rpt_family="MER4D"
complement(14853..15162)
/rpt_family="AluY"
complement(15163..15344)
/rpt_family="MER4D"
complement(15345..15845)
/rpt_family="L1MCS"
16100..16120
/rpt_family="GC-rich"
16155..16177
/rpt_family="GC-rich"
complement(16624..16919)
/rpt_family="AluJb"
complement(18971..19105)
/rpt_family="FLAM A"
complement(19548..19847)
/rpt_family="AluYa5"
23709..23833
/rpt_family="FLAM C"
24811..24889
/rpt_family="MIR"
25425..25642
/rpt_family="C-rich"
27854..27907
/rpt_family="MIR"
28386..28669
/rpt_family="MER41D"
complement(28670..28964)
/rpt_family="AluSx"
28965..29222
/rpt_family="MER41D"

[illegible]

Gardyna,S., Gord,S., Graham,L., Graham,L., Grand-Pierre,N., Hafez,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Nguyen,C., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Racheupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schuback,R., Seaman,S., Severy,P., Smith,C., Stojanovic,N., Talamas,J., Testaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission
Submitted (10-DEC-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 171569)

Birren,B., Nuebaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukngalter,B., Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A., Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S., Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hafez,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schuback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J., Testaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission
Submitted (14-JAN-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
5 (bases 1 to 171569)

Birren,B., Nuebaum,C., Lander,E., Abouelleil,A., Allen,N., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukngalter,B., Camarata,J., Chang,J., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., MacLean,C., Macdonald,P., Major,J., Manning,J., Meneus,L., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Racheupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schuback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J., Testaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission
Submitted (02-APR-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
6 (bases 1 to 171569)

Birren,B., Nuebaum,C., Lander,E., Abouelleil,A., Allen,N., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukngalter,B., Camarata,J., Chang,J., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., MacLean,C.,

Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Racheupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schuback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbbs,M., Talamas,J., Testaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission
Submitted (20-MAY-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On May 20, 2003 this sequence version replaced gi:29469601.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L28146
Center clone name: 516_M_14

FEATURES
source
Location/Qualifiers
1..171569
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="17"
/map="17"
/clone="RP13-516M14"
/clone_lib="RPC1-13 Human Female BAC"
complement(1..74)
/rpt_family="Alu"
complement(520..1125)
/rpt_family="PABL_A"
1768..2096
/rpt_family="AluX"
complement(2186..2397)
/rpt_family="MER66A"
complement(2403..3174)
/rpt_family="MER83A-int"
complement(3768..4157)
/rpt_family="LTR38B"
complement(4159..4439)
/rpt_family="AluX"
4459..4536
/rpt_family="HUESR-P3"
4537..4831
/rpt_family="AluY"
4832..5078
/rpt_family="HUESR-P3"
5079..5377
/rpt_family="AluY"
5378..5490
/rpt_family="HUESR-P3"
5491..5797
/rpt_family="G-rich"
5815..5999
/rpt_family="G-rich"
5925..6094
/rpt_family="(CAGGG)n"
complement(6626..7181)
/rpt_family="MER41E"
complement(7371..7679)
/rpt_family="AluJb"
7716..8020
/rpt_family="AluSg"
8322..8480
/rpt_family="FAM"


```
PC A61P37/06,
PC A61P37/08, A61P43/00, C07K14/47, C07K16/18, C12N1/15, C12N1/19, PC
C12N1/21,
PC C12N5/10, C12P21/02, C12Q1/02, C12Q1/68, G01N33/68// (C12P21/02, PC
C12R1/19),
PC C12N15/00, A61K37/02, C12N5/00
CC 50 human secreted proteins
FH Key Location/Qualifiers
FT source 1..1464
FT source /organism='Homo sapiens (human)'.
FEATURES
    source
        Location/Qualifiers
            1..1464
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
ORIGIN
    Query Match 43.1%; Score 938.6; DB 6; Length 1464;
    Best Local Similarity 95.7%; Pred. No. 1.2e-188;
    Matches 1040; Conservative 0; Mismatches 39; Indels 8; Gaps 7;
Qy 881 AAGGTGCGAGCCCTCAGAGCGGGAGCCAGAGGCGCTGAGCAGAGCCTCGGCTGAAGCTG 940
Db |
Qy 111 AGCGCGCGGGACTCAGAGCGGGAGCCAGAGGCGCTGGCACTGGTGTCAAACCCCTCACCA 170
Db |
Qy 941 TGGACCCAGACTCCGAGCCCAACCCAAAGGCGCTGGCACTGGTGTCAAACCCCTCACCA 1000
Db |
Qy 171 TGGACCCAGACTCCGAGCCCAACCCAAAGGCGCTGGCACTGGTGTCAAACCCCTCACCA 230
Qy 1001 CTTGAGCGCTGGAGCTGTGTCCTCCCAACCCCTGTTTCATATGCGGAGAGCCATA 1060
Db |
Qy 231 CTTGAGCGCTGGAGCTGTGTCCTCCCAACCCCTGTTTCATATGCGGAGAGCCATA 289
Qy 1061 GCGCGCTGCAAGCGAGAGGACACAGAGAGCGAGCCCTGAGTCCGACCTTGGGTGCG 1120
Db |
Qy 230 GCGCGCTGCAAGCGAGAGGACACAGAGAGCGAGCCCTGAGTCCGACCTTGGGTGCG 349
Qy 1121 GGGCGCTGGGTCTCTCGTCCCAACCGAGGGCACAGACACCGGCTTGTGTGGCAGGCTGG 1180
Db |
Qy 350 GGGCGCTGGGTCTCTCGTCCCAACCGAGGGCACAGACACCGGCTTGTGTGGCAGGCTGG 409
Qy 1181 GCCTCTGTGTACCCACTCTGGGTGGGTGCGTGCAGACCTTCCCTCCACCGCCCGAGGCTTT 1240
Db |
Qy 410 GCCTCTGTGTACCCACTCTGGGTGGGTGCGTGCAGACCTTCCCTCCACCGCCCGAGGCTTT 469
Qy 1241 CCAAGCTCTGCTCTCTAGTTTCCAAATGGAACCACTCACTCCGAGACACCGGACTT 1300
Db |
Qy 470 CCAAGCTCTGCTCTCTAGTTTCCAAATGGAACCACTCACTCCGAGACACCGGACTT 529
Qy 1301 ACCAGGACGATGCCCTCTGCTCTGCTCTATCAAAACCCACAGACCCGAGCTCCCTTTCT 1360
Db |
Qy 530 ACCAGGACGATGCCCTCTGCTCTGCTCTATCAAAACCCACAGACCCGAGCTCCCTTTCT 589
Qy 1361 GCCACCCAGGCTGTGTCGGGCCCAAGGTGTGGGGTTCGCTCTCTCACTCCGAGGGCTCC 1420
Db |
Qy 590 GCCACCCAGGCTGTGTCGGGCCCAAGGTGTG -GGGTCCGCTCTCTCACTCCGAGGGCTCC 648
Qy 1421 GGGCCCAAGTGAAGGGGGCCCTGCGGAGCCTCAGACACCTGGAGTTCAGGGC -TGGG 1478
Db |
Qy 649 GGGCCCAAGTGAAGGGGGCCCTGCGGAGCCTCAGACACCTCCAGTTCAGGGCTGTGGG 708
Qy 1479 GGGGCTTTGGCACAATACCTGTGCTGCTGCTATGACAGGCTTTGGGGGCCCTTCCGCGGC 1538
Db |
Qy 709 GGGGCTTTGGCACAATACCTGTGCTGCTGCTATGACAGGCTTTGGGGGCCCTTCCGCGGC 768
Qy 1539 AGCCCCGGGGGGCGAGGTAGGCTGTGGGGGCTTAGAGGCTGGGATGCTCTGTGCCCCAC 1598
Db |
Qy 769 AGCCCCGGGGGGCGAGGTAGGCTG -GGGGGCTTAGAGGCTGGGATGCTCTGTGCCCCAC 827
Qy 1599 GCGCAGGGGGGCAAGCGAGCGCGGCTGGAGGCGCGCGGCTCGGCTCGGGGTGGGGGT 1658
Db |
Qy 828 GCGCAGGGGGGCG -AGCGCAGGGCGGGGTGGAGGCGCGCGGCTCGGGCTGGGGGT 886
```

```
Qy 1659 CAGGTGAGCGCTG -CCTCGGGGCTGGTGGCGATCCCTCAGTCCCTCGGCCACCGGGG 1717
Db |
Qy 887 CAGGTGAGCGCTGCTCCCTCGGGGCTGGTGGCGATCCCTCAGTCCCTCGGCCACCGGGG 946
Qy 1718 GTGCGTCCCTCGTGGCCACCGCAGCTGGGAGCCTTTTGGACCCAGATCTGTTATGCT 1777
Db |
Qy 947 GTGCGTCCCTCGTGGCCACCGCAGCTGGGAGCCTTTTGGACCCAGATCTGTTATGCT 1006
Qy 1778 TTTGCTTTGCTCAGTGGCGGGGCTTTTGTATGTTCTCATCTGAT -GGGGTGGAAAAA 1836
Db |
Qy 1007 TTTGCTTTGCTCAGTGGCGGGGCTTTTGTATGTTCTCATCTGATGGGGTGGAAAAA 1066
Qy 1837 TCACCGGAATCCCTTTCAGTCTTTTGAAGAGTTCCATGCTCGAATATCTGAAATGA 1896
Db |
Qy 1067 TCACCGGAATCCCTTTCAGTCTTTTGAAGAGTTCCATGCTCGAATATCTGAAATGA 1126
Qy 1897 AGAAACAAACCGACTCAAAACCTCCCAAGTAGCTCCAAATGCAATTTTAAATGGAAA 1956
Db |
Qy 1127 AGAAACAAACCGACTCAAAACCTCCCAAGTAGCTCCAAATGCAATTTTAAATGGAAA 1186
Qy 1957 ACAAAAAA 1963
Db |
Qy 1187 AAAAAA 1193
RESULT 13
BD130223
LOCUS BD130223 1064 bp DNA linear PAT 18-SEP-2002
DEFINITION Human signal peptide-containing protein.
ACCESSION BD130223
VERSION BD130223.1 GI:23225168
KEYWORDS JP 2002500009-A/6.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Lai, P., Hillman, J.L., Corley, N.C., Guegler, K.J., Baughn, M.R.,
Sather, S.K. and Shah, P.
TITLE Human signal peptide-containing protein
JOURNAL Patent: JP 2002500009-A 6 08-JAN-2002;
INCYTE PHARMACEUTICALS INC
COMMENT OS Homo sapiens (human)
PS JP 2002500009-A/6
PD 08-JAN-2002
PF 22-DEC-1998 JP 2000526637
PR 31-DEC-1997 US 09/002485
PI PREETI LAL, JENNIFER L HILLMAN, NEIL C CORLEY, KARL J GUEGLER PI
, MARIAH R BAUGHN,
PI SUSAN K SATHER, PURVI SHAH
PC C12N15/09, A61K31/711, A61K35/76, A61K38/00, A61K39/395, A61K39/395, PC
A61K45/00,
PC A61K48/00, A61P31/00, A61P35/00, A61P37/02, A61P37/08, C07K14/47,
C07K16/18,
PC C12N1/00, C12N1/15, C12N1/19, C12N1/21, C12N5/10, C12P21/02, C12Q1/
68,
PC G01N33/68//C12P21/08 C12N15/00 A61K37/02, C12N5/00 CC 1747327
FH Key Location/Qualifiers
FT source 1..1064
FT source /organism='Homo sapiens (human)'.
FEATURES
    source
        Location/Qualifiers
            1..1064
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
ORIGIN
    Query Match 41.9%; Score 912.8; DB 6; Length 1064;
    Best Local Similarity 96.1%; Pred. No. 3.7e-183;
    Matches 957; Conservative 0; Mismatches 37; Indels 2; Gaps 2;
Qy 199 TTCCTGGGGCTCCGGGGCGCGAGAGCTGCATCCAGAGGAGCGCTCCAGAGCGGAC 258
```

[illegible]

RESULT 14
BT008191
LOCUS
DEFINITION
BT008191 linear mRNA 747 bp
Synthetic construct Homo sapiens secreted and transmembrane 1 mRNA, partial cds.

```

BT008191
BT008191.1 GI:30585220
FLI CDNA.
synthetic construct
synthetic construct
other sequences; artificial sequences.
1 (bases 1 to 747)
Kalinine,N., Chen,X., Rolfs,A., Halleck,A., Hines,L., Eisenstein,S.,
Koundinya,M., Raphael,J., Moreira,D., Kelley,T., LaBaer,J., Lin,Y.,
Phelan,M. and Farmer,A.
Cloning of human full-length CDSs in BD Creator(TM) System Donor
vector
Unpublished
2 (bases 1 to 747)
Kalinine,N., Chen,X., Rolfs,A., Halleck,A., Hines,L., Eisenstein,S.,
Koundinya,M., Raphael,J., Moreira,D., Kelley,T., LaBaer,J., Lin,Y.,
Phelan,M. and Farmer,A.
Direct Submission
Submitted (13-MAY-2003) BD Biosciences Clontech, 1020 East Meadow
Circle, Palo Alto, CA 94303, USA
This CDS clone is a part of a collection of human full length
expression clones generated by BD Biosciences Clontech and the
Harvard Institute of Proteomics. Each CDS has been cloned in two
forms: with and without stop-codon (to allow fusion with C-terminal
tag). The CDS has been directionally cloned using BD In-Fusion(TM)
cloning system between the SalI and HindIII sites of the pDNR-DUAL
vector. Additional sequences in the clone: 'ACC' after SalI site
and before 'ATG' to provide Kozak consensus sequence; 'GG' after
last codon and before HindIII site to maintain reading frame.
Clone distribution: http://bioinfo.clontech.com/orfclones.
Location/Qualifiers
1. .747
/organism="synthetic construct"
/mol_type="mRNA"
/db_xref="taxon:32630"
/clone="GH0075011.0"
/clone_lib="BD Creator (TM) CDS Library derived from MGC
collection"
/lab_host="DH5alpha Tl resistant"
/notes="Vector: pDNR-Dual"
1. .>747
/notes="Mutations: 746:Stop->Leu"
/codon_start=1
/transl_table=11
/product="Homo sapiens sapiens secreted and transmembrane 1"
/protein_id="AAP36883.1"
/db_xref="GI:30585221"
/translation="MQTCFLAIFGHCQSALFLFIAALSAGNEGWDSPICTEGVGS
VSGCNTVMSCNTSNISVRIKLRHAGQSALFNEVAFGYSRDQQLQVQGGVAQL
VIKARSHUAGLRYMHLVGHQRNFFTLFLEPKMFAALRAGAQQLSRASALMTPTDSEPTP
VALVWFAYRCRQQORREKLFLEPKMFAALRAGAQQLSRASALMTPTDSEPTP
RPLALVFKPSPLGALSHLSPQLFFPYAADPL"
tch 33.6%; Score 732.8; DB 12; Length 747;
al Similarity 99.6%; Pred. No. 6.3e-145;
745; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
314 ATGCAGACCTGCCCTCGCATTCCTGGCAGCGTTTCCAGGCCCTTGGAGCCCTCTG 373
1 ATGCAGACCTGCCCTCGCATTCCTGGCAGCGTTTCCAGGCCCTTGGAGCCCTCTG 60
374 TTTTGGCTGCTCTCTGATGCTCAGAACTGAAGCTGGAGACGCCCATCTGCACGAG 433
61 TTTTGGCTGCTCTCTGATGCTCAGAACTGAAGCTGGAGACGCCCATCTGCACGAG 120
434 GGCGTAGTCTCTGTGCTTTGGGGCGAGAACCGGTCATGTCCTGCAACATCTCCAGGCC 493
121 GGCGTAGTCTCTGTGCTTTGGGGCGAGAACCGGTCATGTCCTGCAACATCTCCAGGCC 180
494 TTCTCCCATGTCAATCAATGCTGGTCCGACCGGGCAGGAGAGGCCCATCTTCAATGAG 553
181 TTCTCCCATGTCAATCAATGCTGGTCCGACCGGGCAGGAGAGGCCCATCTTCAATGAG 240

```

```
QY 554 GTGGCTCAGGTACTTCTCCGGGACGGCTCGCAGCTCCAGCTTCAGGAGGCGGTGGCA 613
Db |||||
QY 241 GTGGCTCAGGTACTTCTCCGGGACGGCTCGCAGCTCCAGGTTCAAGGAGGCGGTGGCA 300
Db |||||
QY 614 CAGCTGGTGATCAAAAGGCGCGGACCTCCCATGCTGGGCTACATGTGGCACTCTGTG 673
Db |||||
QY 301 CAGCTGGTGATCAAAAGGCGCGGACCTCCCATGCTGGGCTACATGTGGCACTCTGTG 360
Db |||||
QY 674 GGACACCCAGAGAAATAACAGACAAGTCAGCTGGAGTTTCAAGTGCAGAACCCAGTCC 733
Db |||||
QY 361 GGACACCCAGAGAAATAACAGACAAGTCAGCTGGAGTTTCAAGTGCAGAACCCAGTCC 420
Db |||||
QY 734 GGCCTGACACTGGGTTCTGGGCTGTGCCAGCGGTGTCACCTGCTCTTCACTCTCTTG 793
Db |||||
QY 421 GGCCTGACACTGGGTTCTGGGCTGTGCCAGCGGTGTCACCTGCTCTTCACTCTCTTG 480
Db |||||
QY 794 GTCGCTCTGGTATGTCCTGTTACAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAG 853
Db |||||
QY 481 GTCGCTCTGGTATGTCCTGTTGACAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAG 540
Db |||||
QY 854 AAGTCTTCTCTCTAGAACCCAGATGAAGTGCAGGCTCGAGGCGGAGCCAGCAG 913
Db |||||
QY 541 AAGTCTTCTCTAGAACCCAGATGAAGTGCAGGCTCGAGGCGGAGCCAGCAG 600
Db |||||
QY 914 GGCCTGACAGACTTCCGCTGAACTGTGGACCCAGACTCCGAGCCACCCAGGCGCG 973
Db |||||
QY 601 GGCCTGACAGACTTCCGCTGAACTGTGGACCCAGACTCCGAGCCACCCAGGCGCG 660
Db |||||
QY 974 CTGGCACTGGTGTTCACACCTTCCAGTGTGGAGCTGAGCTGCTGCTGCTGCTGCTGCTGCT 1033
Db |||||
QY 661 CTGGCACTGGTGTTCACACCTTCCAGTGTGGAGCTGAGCTGCTGCTGCTGCTGCTGCTGCT 719
Db |||||
QY 1034 CTGTTTTCATATGCGCAGACCCATAG 1061
Db |||||
QY 720 CTGTTTTCATATGCGCAGACCCATAG 747
Db |||||

RESULT 15
AC084737 166458 bp DNA linear HTG 11-NOV-2000
LOCUS Homo sapiens chromosome UL clone RP13-516M14, WORKING DRAFT
DEFINITION AC084737
ACCESSION AC084737
VERSION AC084737.2 GI:11138191
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Waterston,R.H.
The sequence of Homo sapiens clone
Unpublished
REFERENCE 2 (bases 1 to 166458)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (10-NOV-2000) Genome Sequencing Center, Washington
University School of Medicine, 444 Forest Park Parkway, St. Louis,
MO 63108, USA
COMMENT On Nov 11, 2000 this sequence version replaced gi:11138191.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H.FH0516M14
----- Summary Statistics -----
Sequencing vector: M13; 100%
Sequencing vector: plasmid; 0%
Chemistry: Dye-primer ET; 100% of reads
Chemistry: Dye-terminator Big Dye; 0% of reads
```

FEATURES
source

```
Assembly program: Phrap; version 0.990319
Consensus quality: 147705 bases at least Q40
Consensus quality: 155070 bases at least Q30
Consensus quality: 158378 bases at least Q20
Insert size: 172000; agarose-fp
Insert size: 164058; sum-of-contigs
Quality coverage: 3.66 in Q20 bases; agarose-fp
Quality coverage: 3.90 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 25 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1
* 1192: contig of 1192 bp in length
* 1193: gap of unknown length
* 1293: contig of 1538 bp in length
* 2831: gap of unknown length
* 2931: contig of 1555 bp in length
* 4485: gap of unknown length
* 4585: gap of unknown length
* 7134: contig of 2549 bp in length
* 7135: gap of unknown length
* 7235: contig of 2166 bp in length
* 9401: gap of unknown length
* 9501: contig of 1889 bp in length
* 11389: gap of unknown length
* 11390: contig of 2401 bp in length
* 11490: gap of unknown length
* 13891: contig of 2796 bp in length
* 13991: gap of unknown length
* 16787: gap of unknown length
* 16886: gap of unknown length
* 16887: gap of unknown length
* 20617: gap of unknown length
* 20618: gap of unknown length
* 20717: contig of 2814 bp in length
* 23531: gap of unknown length
* 26912: contig of 3282 bp in length
* 26913: gap of unknown length
* 27013: contig of 2477 bp in length
* 29489: gap of unknown length
* 29490: contig of 2939 bp in length
* 29590: gap of unknown length
* 32529: contig of 5847 bp in length
* 32629: gap of unknown length
* 38476: gap of unknown length
* 38575: gap of unknown length
* 38576: gap of unknown length
* 43099: gap of unknown length
* 43199: contig of 3605 bp in length
* 46803: gap of unknown length
* 46804: contig of 4858 bp in length
* 46904: gap of unknown length
* 51762: contig of 5502 bp in length
* 51862: gap of unknown length
* 57363: gap of unknown length
* 57463: gap of unknown length
* 57464: contig of 7758 bp in length
* 65222: gap of unknown length
* 65322: contig of 7050 bp in length
* 72372: gap of unknown length
* 72471: contig of 8220 bp in length
* 72472: gap of unknown length
* 80691: contig of 8220 bp in length
* 80692: gap of unknown length
* 80791: gap of unknown length
* 80792: contig of 7827 bp in length
* 86119: gap of unknown length
* 86119: contig of 20942 bp in length
* 109661: gap of unknown length
* 109661: contig of 20942 bp in length
* 109761: gap of unknown length
* 109761: contig of 22996 bp in length
* 132757: gap of unknown length
* 132757: contig of 33602 bp in length.
* 132857: Location/Qualifiers
* 1.166458
* /organism="Homo sapiens"
* /mol_type="genomic DNA"
* /db_xref="taxon:9606"
* /chromosome="UL"
* /clone="RP13-516M14"
```

Search completed: March 20, 2005, 04:49:43
Job time : 9545 secs

GenCore version 5.1.6

Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 19, 2005, 23:11:41 ; Search time 1185 Seconds
(without alignments)
10890.326 Million cell updates/sec

Title: US-10-080-522-2

Perfect score: 2180

Sequence: 1 ATTCTGCTTCTTTAGCGT.....TTAAATAAGGTTTCGTCGC 2180

Scoring table: IDENTITY NUC

Gapop 10_0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

1: Geneseqn1980s:*

2: Geneseqn1990s:*

3: Geneseqn2000s:*

4: Geneseqn2001as:*

5: Geneseqn2001bs:*

6: Geneseqn2002as:*

7: Geneseqn2002bs:*

8: Geneseqn2003as:*

9: Geneseqn2003bs:*

10: Geneseqn2003cs:*

11: Geneseqn2003ds:*

12: Geneseqn2004as:*

13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2168.8	99.5	2180	2	AAV32446 K12 CDNA
2	1951	89.5	2000	4	AAC88152 Human K12
3	1951	89.5	2000	6	ABL65927 Lung canc
4	1951	89.5	2000	6	ABK84431 Human CDN
5	1951	89.5	2000	12	ADJ74845 Marker ge
6	1951	89.5	2000	13	ADR52746 Drug ther
7	1951	89.5	2000	13	ADP24663 PRO poly
8	1751	80.3	1985	12	ADQ86642 Human tum
9	1751	80.3	1985	13	ADQ85525 Human tum
10	938.6	43.1	1464	3	AAC59772 Human sec
11	912.8	41.9	1064	2	AAX82081 Human SIG
12	692.4	31.8	958	10	ADB36352 Human imm
13	673.8	30.9	942	10	ADB36353 Human imm
14	636.4	29.2	930	9	AAL60549 Human org
15	512.8	23.5	869	9	AAL60550 Human org
16	506.2	23.2	592	10	ADF81837 Leukaemia
17	392.8	18.0	763	9	AAL60548 Human org
18	363.8	16.7	401	10	ADF80286 Leukaemia
19	339.4	15.6	765	9	AAL60551 Human org
20	222	10.2	3026	12	ADM18621 Human chr

c

21	197	9.0	212	2	AAT26356	Aat26356 Human gen	
22	157	7.2	1076	12	ADJ75761	Adj75761 Marker ge	
23	155.4	7.1	1112	4	AAC88154	Aac88154 Mouse K12	
24	155.4	7.1	1112	12	ADJ75760	Adj75760 Marker ge	
25	74.8	3.4	79	4	AAC89734	Aac89734 Human gas	
26	70	3.2	90	4	AAC89735	Aac89735 Human gas	
27	60	2.8	60	6	ABN38487	Abn38487 Human spl	
c	28	56.8	2.6	1127	3	AAO2477	Aao2477 Human col
29	56	2.6	2000	8	ADA71938	Ada71938 Rice gene	
30	54	2.5	100301	6	ABQ88176	Abq88176 Human ost	
31	54	2.5	100301	12	ADQ20754	Adq20754 Human sof	
32	53.2	2.4	114955	2	AAX53491	Aax53491 Human ade	
c	33	51.4	2.4	2000	8	ADA71938	Ada71938 Rice gene
c	34	51.2	2.3	5452	10	ADC86736	Adc86736 Human GPC
c	35	51	2.3	114955	2	AAX53491	Aax53491 Human ade
c	36	49.8	2.3	660	5	AAS68601	Aas68601 DNA encod
c	37	49.6	2.3	1679	8	ABZ78139	Abz78139 Human can
c	38	49.4	2.3	4667	12	ADQ22939	Adq22939 Human sof
c	39	49.4	2.3	5452	10	ADC86736	Adc86736 Human GPC
c	40	48.4	2.2	3897	12	ADJ78312	Adj78312 Human per
c	41	48.4	2.2	22210	12	ADJ78311	Adj78311 Human per
c	42	48	2.2	647	4	AAH06718	Aah06718 Human CDN
c	43	48	2.2	2437	4	AAH18557	Aah18557 Human CDN
c	44	47.6	2.2	8459	4	AAC89557	Aac89557 Human his
c	45	47.6	2.2	8459	8	ACC47997	Acc47997 Human his

ALIGNMENTS

RESULT 1

AAV32446

ID AAV32446 standard; cDNA; 2180 BP.

XX AC AAV32446;

XX DT 26-OCT-1998 (first entry)

XX DE K12 CDNA sequence.

XX KW DNA probe; CD7 HSI DNase hypersensitive site; mRNA northern blot;

KW human erythroleukemic; HEL; K562 cell line; clone; breast cancer;

KW ovarian cancer; malignant; tumour; ss.

XX OS Homo sapiens.

XX FH Key

XX CDS

XX Location/Qualifiers

XX 314..1061

XX /*tag= a

XX /transl_except= (pos:1028..1031, aa:Gln)

XX /note= "this codon has an apparent 1 nucleotide

XX insertion, which alters the reading frame"

XX

XX WO9822502-A1.

XX

XX 28-MAY-1998.

XX

XX 24-NOV-1997; 97WO-US021517.

XX

XX 22-NOV-1996; 96US-00755559.

XX

XX (UYDU-) UNIV DUKE.

XX

XX Kaufman RE, Slentz-Kesler KA;

XX

XX WPI; 1998-312415/27.

XX P-PSDB; AAW48811.

XX

XX New isolated K12 protein gene - which is over expressed in certain

XX neoplastic cells, used to develop products for tumour detection and

XX treatment.

XX

XX Claim 8; Fig 2A-2B; 44pp; English.

XX PS

XX This present sequence represents the K12 cDNA, which has been located on
CC chromosome 17q25. To obtain the cDNA sequence a 500 bp DNA probe, which
CC can be located just upstream of the CD7 Hs1 DNase hypersensitive site,
CC was used against a mRNA northern blot. From this a 1.8kb transcript was
CC detected in the human erythroleukemic cell line HEL. The probe was then
CC used to screen a human erythroleukemic cell line K562 cDNA library, from
CC which several clones were identified and isolated that constituted a
CC 1.8kb cDNA. This cDNA was designated K12 and was found to have a single
CC open reading frame as well as being in the same orientation as CD7. The
CC K12 gene was found to be expressed in both breast and ovarian cancer
CC cells at a much higher level than any other malignant or normal tissue
CC that was examined, thus enabling the K12 to be a useful protein in tumour
CC detection and treatment

XX
CC
COMM-000000 0180 BB. 391 A. 730 C. 639 G: 420 T: 0 U: 0 Other;

Query Match	99.5%	Score 2168.8	DB 2	Length 2180
Best Local Similarity	99.7%	Pred. NO. 0		
Matches 2173	Conservative	0	Mismatches 7	Indels 0
				Gaps 0

QY	1	ATTCTGCTTCTTTAGCGTGAACGCGGTGCGTCTCCGTGAAATAATAAATTCAC	60
DB	1	ATTCTGCTTCTTTAGCGTGAACGCGGTGCGTCTCCGTGAAATAATAAATTCAC	60
QY	61	CGTCACGCTTGTGTGAACGCGGTGTGTTCCGAAACTTGGAGGCTTCCCGTAACCCAG	120
DB	61	CGTCACGCTTGTGTGAACGCGGTGTGTTCCGAAACTTGGAGGCTTCCCGTAACCCAG	120
QY	121	CTCCTTCTCATCTCGGAGGTGGGTCCGCGGGTCCGCGCTTCCCTCCCTGGCGCCCTC	180
DB	121	CTCCTTCTCATCTCGGAGGTGGGTCCGCGGGTCCGCGCTTCCCTCCCTGGCGCCCTC	180
QY	181	CTCTCTGCTCTTTTCAATTTTCTTGGGGTCTCGGGCGCGGAGAAAGCTGCATCCGAGGA	240
DB	181	CTCTCTGCTCTTTTCAATTTTCTTGGGGTCTCGGGCGCGGAGAAAGCTGCATCCGAGGA	240
QY	241	GGCGCTCAGGAGCGGACCGGGAGTGTTCGAAGGCCAGTGAACAAGCAACACGAGGCCCA	300
DB	241	GGCGCTCAGGAGCGGACCGGGAGTGTTCGAAGGCCAGTGAACAAGCAACACGAGGCCCA	300
QY	301	AGTCCCAACACGCAATCAGACACTGCCCTCGCATTTCCCTGTGCCACAGTTTCCAGGCCCT	360
DB	301	AGTCCCAACACGCAATCAGACACTGCCCTCGCATTTCCCTGTGCCACAGTTTCCAGGCCCT	360
QY	361	TGGGACCTCTCTGTTTTTGGCTGCTCTTGAAGTGTCTCAGATGAAGGCTGGGACAGGCC	420
DB	361	TGGGACCTCTCTCTGTTTTTGGCTGCTCTTGAAGTGTCTCAGATGAAGGCTGGGACAGGCC	420
QY	421	CATCTGCACAGGGGTAGTCTCTGTCTTGGGGCGAGAAACAAGTCAATGTCTTGCA	480
DB	421	CATCTGCACAGGGGTAGTCTCTGTCTTGGGGCGAGAAACAAGTCAATGTCTTGCA	480
QY	481	CATCTCAAGCCCTTCTCCATGTCACATCAAGCTCGCTGCCACCGGCGAGAGAGCGC	540
DB	481	CATCTCAAGCCCTTCTCCATGTCACATCAAGCTCGCTGCCACCGGCGAGAGAGCGC	540
QY	541	CATCTTCAATGAGGTGCTCCAGGCTACTTCTCCGGGACGGCTGCGAAGCTCAGGTTCA	600
DB	541	CATCTTCAATGAGGTGCTCCAGGCTACTTCTCCGGGACGGCTGCGAAGCTCAGGTTCA	600
QY	601	GGGAGGCGTGGGACAGCTGGTGATCAAAAGCGCCGGGACTCCCATGCTGGGCTGTACAT	660
DB	601	GGGAGGCGTGGGACAGCTGGTGATCAAAAGCGCCGGGACTCCCATGCTGGGCTGTACAT	660
QY	661	GTGCGACCTCGTGGGACACGAGAGAAATAACAGAAAGTCAAGCTGGAGGTTTCAGGTGC	720
DB	661	GTGCGACCTCGTGGGACACGAGAGAAATAACAGAAAGTCAAGCTGGAGGTTTCAGGTGC	720
QY	721	AGAAACCCAGTCCCGCCCTGACACTGGGTTCTGGCTGTGCCAGCGGTGTCACTGCTGT	780
DB	721	AGAAACCCAGTCCCGCCCTGACACTGGGTTCTGGCTGTGCCAGCGGTGTCACTGCTGT	780

Db 1861 TTTGAAAAGTTCATGACTCGAATATCTGAATGAAGAAAACAAACCGACTCACAAACC 1920
Qy 1921 TCCAGTAGTCTCAATGCAATTTTAAATGGAACCAAAATCTGAAAGAAAGCTCTT 1980
Db 1921 TCCAGTAGTCTCAATGCAATTTTAAATGGAACCAAAATCTGAAAGAAAGCTCTT 1980
Qy 1981 TAGTGGCTTTAAGCCCAAAACGCTCCCTAAGGCGTCTCGAGATGAAGACGGGGGAGC 2040
Db 1981 TAGTGGCTTTAAGCCCAAAACGCTCCCTAAGGCGTCTCGAGATGAAGACGGGGGAGC 2040
Qy 2041 CCCAGCCAGGTGAGACCCGAGACGGCGGCGCGCTGACCGAGGCTCGACAG 2100
Db 2041 CCCAGCCAGGTGAGACCCGAGACGGCGGCGCGCTGACCGAGGCTCGACAG 2100
Qy 2101 CCGGCGCGCTGAGGCTGGGCGGAGCCAGGCTCAAGAGGGCGCTTTGTCTCGGG 2160
Db 2101 CCGGCGCGCTGAGGCTGGGCGGAGCCAGGCTCAAGAGGGCGCTTTGTCTCGGG 2160
Qy 2161 TTAATAAAGTTCCGTCG 2180
Db 2161 TTAATAAAGTTCCGTCG 2180

RESULT 2

ID AAC88152 standard; cDNA; 2000 BP.
AC AAC88152;
XX
DT 13-MAR-2001 (first entry)
DE Human K12 nucleotide sequence SEQ ID NO:3.
XX
KW Human; CD7; K12; cognate ligand; cluster of differentiation; cancer;
KW identification; inhibiting T cell proliferation; HIV; infection;
KW activating natural killer cell proliferation; leukaemia; lymphoma;
KW sepsis; graft versus host disease; autoimmune disease; arthritis;
KW multiple sclerosis; rheumatoid arthritis; psoriatic arthritis; lupus;
KW scleroderma; psoriasis; atopic dermatitis; type I diabetes mellitus;
KW Hashimoto's thyroiditis; pernicious anaemia; Addison's disease; uveitis;
KW myasthenia gravis; psoriasis; Guillain-Barre Syndrome; Grave's disease;
KW systemic lupus erythematosus; dermatomyositis; asthma; eczema;
KW atopic dermatitis; contact dermatitis; eczematous dermatitis;
KW seborrheic dermatitis; rhinitis; ss.
XX
OS Homo sapiens.
XX
PN WO200073333-A2.
XX
PD 07-DEC-2000.
XX
PF 26-MAY-2000; 2000WO-US014612.
XX
PR 28-MAY-1999; 99US-0136450P.
XX
PA (IMM V) IMMUNEX CORP.
XX
PI Lyman SD, Fanslow WC;
XX
DR WPI; 2001-061511/07.
XX
DR P-PSDB; AAB36658.
XX
PT Stimulating intracellular signaling of CD7 comprises contacting a cell
PT expressing CD7 with recombinant K12 protein, the cognate ligand of CD7,
PT to inhibit T cell proliferation and/or activate natural killer cell
PT proliferation.
XX
PS Disclosure; Page 36-38; 42pp; English.
XX
CC The present invention describes a method for stimulating (S) the
CC intracellular signalling of CD (Cluster of differentiation) 7 comprising
CC contacting a cell that expresses CD7 with a recombinant K12 protein (I),

CC the cognate ligand of CD7. (S) is useful for inhibiting T cell
CC proliferation and/or activating NK (natural killer) cell proliferation
CC and/or inducing NK toxicity in a mammal which involves administration of
CC K12 protein. It is also used for treating HIV-1 infection, cancer (T cell
CC leukaemia, acute lymphocytic leukaemia, cutaneous T cell lymphoma),
CC bacterial and viral infections, mediated by CD7. In the case of treating
CC T cell leukaemia the soluble K12 protein is covalently attached to a
CC toxin. A disease mediated by CD7 such as sepsis, graft versus host
CC disease due to transplantation, autoimmune diseases, multiple sclerosis,
CC arthritis, rheumatoid arthritis, psoriatic arthritis, scleroderma, lupus,
CC psoriasis, atopic dermatitis, type I diabetes mellitus, Hashimoto's
CC thyroiditis, pernicious anaemia, Addison's disease, myasthenia gravis,
CC uveitis, psoriasis, Guillain-Barre Syndrome, Grave's disease, systemic
CC lupus erythematosus and dermatomyositis, asthma, eczema, atypical
CC dermatitis, contact dermatitis, other eczematous dermatitides, seborrheic
CC dermatitis, and rhinitis is also treated by administering a K12
CC antagonist (neutrallising antibody). The present sequence encodes the
CC human K12 protein, which is given in the exemplification of the present
CC invention
XX
SQ Sequence 2000 BP; 367 A; 659 C; 602 G; 372 T; 0 U; 0 Other;
Query Match 89.5%; Score 1951; DB 4; Length 2000;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1984; Conservative 0; Mismatches 0; Indels 3; Gaps 3;
Qy 196 ATTTTCTGGGGCTCCGGGGCGGAGAAAGTGCATCCAGAGGAGCGCTCCAGGAGCG 255
Db 1 ATTTTCTGGGGCTCCGGGGCGGAGAAAGTGCATCCAGAGGAGCGCTCCAGGAGCG 60
Qy 256 GACCCGGAGGTGTTTCAAGAGCCAGTGACAAAGGACCAGGGGCCCAAGTCCCAAGCCAT 315
Db 61 GACCCGGAGGTGTTTCAAGAGCCAGTGACAAAGGACCAGGGGCCCAAGTCCCAAGCCAT 120
Qy 316 GCAGACCTGCCCCCTGGCATTCCTTGGCCAGCTTCCAGGCGCTTGGGACCTCTCTGT 375
Db 121 GCAGACCTGCCCCCTGGCATTCCTTGGCCAGCTTCCAGGCGCTTGGGACCTCTCTGT 180
Qy 376 TTTGGCTGCTCTCTGTTGAGTGTCTAGATGCTCAGATGAAGGCTGGGACAGCCCATCTGCACAGAGGG 435
Db 181 TTTGGCTGCTCTCTGTTGAGTGTCTAGATGAAGGCTGGGACAGCCCATCTGCACAGAGGG 240
Qy 436 GGTAGTCTCTGTGTTGTTGGGGCGAGAACACCGTGTATGCTTGCAACATCTTCAAGCGCTT 495
Db 241 GGTAGTCTCTGTGTTGTTGGGGCGAGAACACCGTGTATGCTTGCAACATCTTCAAGCGCTT 300
Qy 496 CTCCCATGTCAACATCAAGCTGCGTCCAGCGGACGAGAGCGCCATCTTCAATGAGGT 555
Db 301 CTCCCATGTCAACATCAAGCTGCGTCCAGCGGACGAGAGCGCCATCTTCAATGAGGT 360
Qy 556 GGCTCCAGGCTACTTCTCCCGGACGCGTGGCAGCTCCAGGTTTCAGGGAGGCGTGGCACA 615
Db 361 GGCTCCAGGCTACTTCTCCCGGACGCGTGGCAGCTCCAGGTTTCAGGGAGGCGTGGCACA 420
Qy 616 GCTGGTGATCAAAAGGCGCCGGGACTCCCATGTGGGTGTATGTGGCACCTCGTGGG 675
Db 421 GCTGGTGATCAAAAGGCGCCGGGACTCCCATGTGGGTGTATGTGGCACCTCGTGGG 480
Qy 676 ACACGAGAGAAATACAGACAAAGTTCAGCTGGAGGTTTCAGGTGCAGAACCCAGTCCGC 735
Db 481 ACACGAGAGAAATACAGACAAAGTTCAGCTGGAGGTTTCAGGTGCAGAACCCAGTCCGC 540
Qy 736 CCCTGACACTGGGTTCTGGCCTGTGCCAGCGGTGTCTACTGTCTTTCATCTCTTGGT 795
Db 541 CCCTGACACTGGGTTCTGGCCTGTGCCAGCGGTGTCTACTGTCTTTCATCTCTTGGT 600
Qy 796 CGCTCTGTCTCATGTTGCGCTTGGTACAGGTGCGCTGTTTCCAGCAACGCCGGGAGAGAA 855
Db 601 CGCTCTGTCTCATGTTGCGCTTGGTACAGGTGCGCTGTTTCCAGCAACGCCGGGAGAGAA 660
Qy 856 GTTCTTCTCTAGAACCCAGATGAAGGTTCGACGCTTCAGAGCGGAGCCAGCAGGG 915
Db 661 GTTCTTCTCTAGAACCCAGATGAAGGTTCGACGCTTCAGAGCGGAGCCAGCAGGG 720

QY	916	CCTGAGCAGAGCCTCCGCTGAACTGTGAGACCCACAGATCCGAGGCCACCCCAAGGCCGCT	975
DB	721	CCTGAGCAGAGCCTCCGCTGAACTGTGAGACCCACAGATCCGAGGCCACCCCAAGGCCGCT	780
QY	976	GGCACCTGGTTCMAACCCCTCACCACTTTGGAGCCCTGGAGCTGTGTCTCCGCCCAACCCCT	1035
DB	781	GGCACTGGTGTTCMAACCCCTCACCACTTTGGAGCCCTGGAGCTGTGTCTCCGCCCAACCCCT	839
QY	1036	TGTTTCCATATGCGCAGACCCATAGCCCTCTGCAAGGCGAGAGGACACAGGAGAGCCA	1095
DB	840	TGTTTCCATATGCGCAGACCCATAGCCCTCTGCAAGGCGAGAGGACACAGGAGAGCCA	899
QY	1096	GCCTGAGTGCAGACCTTGGGTGGGGGCGCTGGGTCTCTCGTCCACCCCGAGAGGGCACA	1155
DB	900	GCCTGAGTGCAGACCTTGGGTGGGGGCGCTGGGTCTCTCGTCCACCCCGAGAGGGCACA	959
QY	1156	GACACCGGCTTGCTTGGCAGGCTGGGCTCTGTGTACCCATCTCTGGGTGCGTGCAGAC	1215
DB	960	GACACCGGCTTGCTTGGCAGGCTGGGCTCTGTGTACCCATCTCTGGGTGCGTGCAGAC	1019
QY	1216	CTTTCCTCCATCCACCCCGCAGCTCTCCAGCTCTGCTTCTCAGTTTCCAAAATGGAAACC	1275
DB	1020	CTTTCCTCCATCCACCCCGCAGCTCTCCAGCTCTGCTTCTCAGTTTCCAAAATGGAAACC	1079
QY	1276	ACCTCACTCTCGCAGCACCCGACTTACCAGGACGCAATGCCCTCCCTCTGCTTCATCAA	1335
DB	1080	ACCTCACTCTCGCAGCACCCGACTTACCAGGACGCAATGCCCTCCCTCTGCTTCATCAA	1139
QY	1336	ACCACAGACCCGCACTCCCTTTCTGCCACCCAGGCTGTGTCGGCCCCCAGGTGTGGGT	1395
DB	1140	ACCACAGACCCGCACTCCCTTTCTGCCACCCAGGCTGTGTCGGCCCCCAGGTGTGGGT	1199
QY	1396	CGCTCTCTCCACCTCCACAGGCTCCGCGCCCAAGTGAGGGGGCCCTGCGGAGCCTCAG	1455
DB	1200	CGCTCTCTCCACCTCCACAGGCTCCGCGCCCAAGTGAGGGGGCCCTGCGGAGCCTCAG	1259
QY	1456	ACACATGGAATTCAGGGCTGGGGGGCCCTTGCAATACCTGTCCCTTGGCTATGAGCA	1515
DB	1260	ACACATGGAATTCAGGGCTGGGGGGCCCTTGCAATACCTGTCCCTTGGCTATGAGCA	1319
QY	1516	GGCTTTGGGGGCCCTTCGCGGCGACGCCCGGGGCGGAGTAGGGTCTGGGGGCTTAGAG	1575
DB	1320	GGCTTTGGGGGCCCTTCGCGGCGACGCCCGGGGCGGAGTAGGGTCTGGGGGCTTAGAG	1379
QY	1576	GCTGGGATGCTCTCTGGCCCCACCGCAGAGGGGCGAGCGCGGCTGGAGGCGGC	1635
DB	1380	GCTGGGATGCTCTCTGGCCCCACCGCAGAGGGGCGAGCGCGGCTGGAGGCGGC	1439
QY	1636	GGCGCGCTCTGGGCTGGGGGGTCAAGTGAACGCTGCTCCGGGGCTGGTTCGCGATCCC	1695
DB	1440	GGCGCGCTCTGGGCTGGGGGGTCAAGTGAACGCTGCTCCGGGGCTGGTTCGCGATCCC	1499
QY	1696	TCAGTCCCTCGGCCACCCGGGGGTGCTCCCTCGTGCACACCGCACTGCGGAGCCTCTT	1755
DB	1500	TCAGTCCCTCGGCCACCCGGGGGTGCTCCCTCGTGCACACCGCACTGCGGAGCCTCTT	1559
QY	1756	TGGAACCCAGATCTGTTCAATGCTTTGTCTGCTCACTCGCGCGGGGCCCTTTGATGCTT	1815
DB	1560	TGGAACCCAGATCTGTTCAATGCTTTGTCTGCTCACTCGCGCGGGGCCCTTTGATGCTT	1619
QY	1816	CATCTGTATGGGTGGAAATACCGGGAATCCCCCTTCAGTCTTTTGAAGAAGTTCCA	1875
DB	1620	CATCTGTATGGGTGGAAATACCGGGAATCCCCCTTCAGTCTTTTGAAGAAGTTCCA	1679
QY	1876	TGACTCGAATATCTGAATGAAGAAACAAACCCGACTCACAACTCCAAAGTAGCTCCAA	1935
DB	1680	TGACTCGAATATCTGAATGAAGAAACAAACCCGACTCACAACTCCAAAGTAGCTCCAA	1739
QY	1936	ATGCAATTTTTAAATTGAATAACAAATCTGTGAAGAAGACGCTTTTAGTGCTTTAGCC	1995
DB	1740	ATGCAATTTTTAAATTGAATAACAAATCTGTGAAGAAGACGCTTTTAGTGCTTTAGCC	1799

PR 02-OCT-2000; 2000US-0237295P.
PR 02-OCT-2000; 2000US-0237316P.
PR 03-OCT-2000; 2000US-0237425P.
PR 03-OCT-2000; 2000US-0237598P.
PR 03-OCT-2000; 2000US-0237604P.
PR 03-OCT-2000; 2000US-0237606P.
PR 03-OCT-2000; 2000US-0237608P.
PR 01-NOV-2000; 2000US-0244867P.
PR 01-NOV-2000; 2000US-0245084P.
XX (AVAL-) AVALON PHARM.
XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX WPI; 2002-188264/24.
XX Screening for anti-neoplastic agent involves exposing cells to a chemical
PT agent to be tested for anti-neoplastic activity, and determining a change
PT in expression of a gene of a signature gene set.
XX Claim 1; SEQ ID NO 4264; 44pp; English.
XX The present invention describes a method (M1) for screening for an anti-
CC neoplastic agent. The method involves exposing cells to a chemical agent
CC to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening an
CC anti-neoplastic agent, and can be used for producing a product which is
CC the data collected with respect to the anti-neoplastic agent as a result
CC of M1, and the data is sufficient to convey the chemical structure and/or
CC properties of the agent. M1 can be used in the treatment of cancer such
CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's
CC tumour
XX
SQ Sequence 2000 BP; 367 A; 659 C; 602 G; 372 T; 0 U; 0 Other;

Query Match 89.5%; Score 1951; DB 6; Length 2000;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1984; Conservative 0; Mismatches 0; Indels 3; Gaps 3;

QY 196 ATTTTCTGGGCTCGGGGCGGAGAGCTGCATCCAGAGAGCGGTCACAGGG 255
DB |||||
QY 256 GACCCGGAGTGTTCAGAGGCACTGACAGGACACAGGGGCCAAGTCCACAGCCAT 315
DB |||||
QY 61 GACCCGGAGTGTTCAGAGGCACTGACAGGACACAGGGGCCAAGTCCACAGCCAT 120
DB |||||
QY 316 CGAGACCTGCCCTCGGCAATCCCTGGCCACGTTTCCAGGCGCTTGGACCCCTCTGT 375
DB |||||
QY 121 CGAGACCTGCCCTCGGCAATCCCTGGCCACGTTTCCAGGCGCTTGGACCCCTCTGT 180
DB |||||
QY 376 TTGTGCTGCTCTTGTAGTGTCTCAGATGAAGGCTGGAGACGCCCATCTGCACAGGG 435
DB |||||
QY 181 TTGTGCTGCTCTTGTAGTGTCTCAGATGAAGGCTGGAGACGCCCATCTGCACAGGG 240
DB |||||
QY 436 GGTAGTCTCTGTGTCTTGGGGGAGAACACCGTCAATGCTTCCAACTCTCCAGCCCTT 495
DB |||||
QY 241 GGTAGTCTCTGTGTCTTGGGGGAGAACACCGTCAATGCTTCCAACTCTCCAGCCCTT 300
DB |||||
QY 496 CTCCCATGTCAACATCAAGCTCGTCCCGGAGGAGCGCCATCTTCAATGAGGT 555
DB |||||
QY 301 CTCCCATGTCAACATCAAGCTCGTCCCGGAGGAGCGCCATCTTCAATGAGGT 360
DB |||||
QY 556 GCTTCCAGGCTACTTCTCCGGAGCGGTGGAGCTCAGGTTTCAAGGAGGCGTGGCA 615
DB |||||

DB 361 GGCTCCAGGCTACTTCTCCCGGAGACGGCTGGCAGCTCCAGGTTTCAAGGAGGCGTGGCACA 420
QY 616 GCTGTGATCAAAAGGCGCCGGGACTCCCATGCTGGGCTGTACATGTGGCAGCTCTGTGGG 675
DB 421 GCTGTGATCAAAAGGCGCCGGGACTCCCATGCTGGGCTGTACATGTGGCAGCTCTGTGGG 480
QY 676 ACACAGAGAAATAACAGACAAAGTCA CGCTGGAGGTTTCAAGGTGAGAAACCCAGTCCGC 735
DB 481 ACACAGAGAAATAACAGACAAAGTCA CGCTGGAGGTTTCAAGGTGAGAAACCCAGTCCGC 540
QY 736 CCCTGACACTGGGTTCTGGCCCTGTGCCAGCGGTGGTCACTGCTGCTTCTCATCTCTTGGT 795
DB 541 CCCTGACACTGGGTTCTGGCCCTGTGCCAGCGGTGGTCACTGCTGCTTCTCATCTCTTGGT 600
QY 796 CGCTCTGCTCATGTTTCGCTCTGTA CAGGTGCGCTGTTTCCAGCAACCGCGGAGAA 855
DB 601 CGCTCTGCTCATGTTTCGCTCTGTA CAGGTGCGCTGTTTCCAGCAACCGCGGAGAA 660
QY 856 GTTCTTCTCTAGAACCCAGATGAAGGTGCGAGCCCTCAGAGCGGAGCCAGAGGG 915
DB 661 GTTCTTCTCTAGAACCCAGATGAAGGTGCGAGCCCTCAGAGCGGAGCCAGAGGG 720
QY 916 CCTGAGCAGAGCTCCGCTGAACCTGTGGAACCCCGAGACTCCGAGGCCACCCCAAGGCCCT 975
DB 721 CCTGAGCAGAGCTCCGCTGAACCTGTGGAACCCCGAGACTCCGAGGCCACCCCAAGGCCCT 780
QY 976 GGCCTGCTGTTTCAAAACCCCTCA CCACTTGGAGCCCTGGAGCTGTGCTCCGCCCAACCT 1035
DB 781 GGCCTGCTGTTTCAAAACCCCTCA CCACTTGGAGCCCTGGAGCTGTGCTGTGCTGTGCT 839
QY 1036 TGTTCCTATATGCGCGAGACCCATAG CCGCTGTGAAGCGAGAGGACACAGGAGGCA 1095
DB 840 TGTTCCTATATGCGCGAGACCCATAG CCGCTGTGAAGCGAGAGGACACAGGAGGCA 899
QY 1096 GCCCTGAGTGGCGACCTTGGGTGGG GCGGCTGGGTCTCTGTCCTCCACCGGAGGCGACA 1155
DB 900 GCCCTGAGTGGCGACCTTGGGTGGG GCGGCTGGGTCTCTGTCCTCCACCGGAGGCGACA 959
QY 1156 GACACCGGCTTGTGTCGAGGCTGGG GCGCTCTGTGTACCCACTCTCGGGTGGTGCAGAC 1215
DB 960 GACACCGGCTTGTGTCGAGGCTGGG GCGCTCTGTGTACCCACTCTCGGGTGGTGCAGAC 1019
QY 1216 CTTTCCCTCTCACCCCTCAGGCTTTC CAAAGCTCTGCTTCTCAGTTTCCAAATGGAAACC 1275
DB 1020 CTTTCCCTCTCACCCCTCAGGCTTTC CAAAGCTCTGCTTCTCAGTTTCCAAATGGAAACC 1079
QY 1276 ACCTCACCTCCGAGCACCAGACTTAC CAGAGCCATGCGCTTCTGCTGCTCCCTCATCAA 1335
DB 1080 ACCTCACCTCCGAGCACCAGACTTAC CAGAGCCATGCGCTTCTGCTGCTCCCTCATCAA 1139
QY 1336 ACCCAGACACCGGACTCCCTTTTCTG CCAACCCAGGCTGGTCCGGCCCGCAGGTGGGGT 1395
DB 1140 ACCCAGACACCGGACTCCCTTTTCTG CCAACCCAGGCTGGTCCGGCCCGCAGGTGGGGT 1199
QY 1396 CGCTCTCTCACTCCAGGGCTCCGCGCC CAAAGTGGAGGGGCCCTTCCCGAGGCTCAG 1455
DB 1200 CGCTCTCTCACTCCAGGGCTCCGCGCC CAAAGTGGAGGGGCCCTTCCCGAGGCTCAG 1259
QY 1456 ACACACTGGAGTTCAGGGCTGGGGGGG GCGCTTGGGCACATACCTGCTTGGCTATGACA 1515
DB 1260 ACACACTGGAGTTCAGGGCTGGGGGG GCGCTTGGGCACATACCTGCTTGGCTATGACA 1319
QY 1516 GGCTTTTGGGGGCCCTTCCGCGCGAGCC CCGGGGGCCGAGGTAGGCTTGGGGGCTTAGAG 1575
DB 1320 GGCTTTTGGGGGCCCTTCCGCGCGAGCC CCGGGGGCCGAGGTAGGCTTGGGGGCTTAGAG 1379
QY 1576 GCTGGAGTGGCTCTCTGGCCCCCAACG CCAAGGGGGCAAGCGAGGCGGCTGGAGGCGGC 1635
DB 1380 GCTGGAGTGGCTCTCTGGCCCCCAACG CCAAGGGGGCAAGCGAGGCGGCTGGAGGCGGC 1439
QY 1636 GCGGCGGCTCGGGCTGGGGGGTCAAGT GAGAGCTGCTTCCGGGGCTGGTCCGCGATCCC 1695
DB 1440 GCGGCGGCTCGGGCTGGGGGGTCAAGT GAGAGCTGCTTCCGGGGCTGGTCCGCGATCCC 1499

QY	1696	TCAGTCCCTCGCCACCCCGGGGTGCGTCCCTCGTGGCCACGACCTCCGAGCCTCTT	1755
Db	1500	TCAGTCCCTCGCCACCCCGGGGTGCGTCCCTCGTGGCCACGACCTCCGAGCCTCTT	1559
QY	1756	TGAGCCAGATCTGTTTCATGCTTTTGTCTTCTGTCACCTGGGGGCTTTTGATGCTT	1815
Db	1560	TGAGCCAGATCTGTTTCATGCTTTTGTCTTCTGTCACCTGGGGGCTTTTGATGCTT	1619
QY	1816	CATCTGTATGGGTGGAATAATCACCGGAATCCCTTCAGTTCTTTTGAAGAAAGTTCCA	1875
Db	1620	CACTGTATGGGTGGAATAATCACCGGAATCCCTTCAGTTCTTTTGAAGAAAGTTCCA	1679
QY	1876	TGACTCGAATATCTGAATGAAGAAAACAAACCGACTCACAACTCCAAAGTAGTCCAA	1935
Db	1680	TGACTCGAATATCTGAATGAAGAAAACAAACCGACTCACAACTCCAAAGTAGTCCAA	1739
QY	1936	ATGCAATTTTAAATGGAACAAATAATCTGAAGAAACGCTTTTAGTGGCTTTAAGCC	1995
Db	1740	ATGCAATTTTAAATGGAACAAATAATCTGAAGAAACGCTTTTAGTGGCTTTAAGCC	1799
QY	1996	CAAAACGTCCTTAAGCGCTCTCGAGATGAAGACGGGGGAG-CCCCAGCCAGGTGGA	2054
Db	1800	CCAAACGTCCTTAAGCGCTCTCGAGATGAAGACGGGGGAGCCCCAGCCAGGTGGA	1859
QY	2055	GACCCGCGAGAGCGCGCGCGCGCGTACCGAGGCTCGCACAGCGCGCGCCCTGAG	2114
Db	1860	GACCCGCGAGAGCGCGCGCGCGTACCGAGGCTCGCACAGCGCGCGCCCTGAG	1919
QY	2115	GGTCCGGCC-GAGCCAGGCTCCAAAGAGGGCGGTTGTGTCTCGGTTAAATAAGGTT	2173
Db	1920	GGTCCGGCCGAGCCAGGCTCCAAAGAGGGCGGTTGTGTCTCGGTTAAATAAGGTT	1979
QY	2174	CGGTCCG 2180	
Db	1980	CGGTCCG 1986	
RESULT 4			
ABK84431			
ID	ABK84431 standard; cDNA; 2000 BP.		
AC	ABK84431;		
XX			
XX	14-AUG-2002 (first entry)		
DT	Human cDNA differentially expressed in granulocytic cells #1002.		
DE			
XX	Human; ss; granulocytic cell; DNA chip; bacterial infection;		
KW	viral infection; parasitic infection; protozoal infection;		
KW	fungal infection; sterile inflammatory disease; psoriasis;		
KW	rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;		
KW	cardiac reperfusion injury; renal reperfusion injury; ARDS;		
KW	adult respiratory distress syndrome; inflammatory bowel disease;		
KW	Crohn's disease; ulcerative colitis; periodontal disease;		
KW	granulocyte activation; chronic inflammation; allergy.		
OS	Homo sapiens.		
XX	WO200228999-A2.		
FN	11-APR-2002.		
PD			
XX	03-OCT-2001; 2001WO-US030821.		
PF			
XX	03-OCT-2000; 2000US-0237189P.		
PR			
XX	(GENE-) GENE LOGIC INC.		
PA			
XX	Beazer-Barclay Y, Weisman SM, Yamaga S, Vockley J;		
FI	WPI; 2002-435328/46.		
XX			
DR			
XX			

Detecting granulocyte activation by detecting differential expression of genes associated with granulocyte activation, which serves as diagnostic markers that is useful for monitoring disease states and drug toxicity.

Claim 1; SEQ ID NO 1002; 114pp; English.

The invention relates to detecting (M1) granulocyte (GC) activation (GCA) by detecting the level of expression of gene(s) (Gs) identified by DNA chip analysis as given in the specification, and comparing the expression level to an expression level in an unactivated GC, where differential expression of Gs is indicative of GCA. Also included are modulating (M2) GA by contacting GC with an agent that alters the expression of at least one gene in Gs; (2) screening (M3) for an agent capable of modulating GCA or an inflammation (especially chronic) in a tissue, an allergic response in a subject, exposure of a subject to a pathogen or sterile inflammatory disease using the gene expression profile; (3) detecting (M4) an inflammation (especially chronic) in a tissue, an allergic response in a subject, exposure of a subject to a pathogen or sterile inflammatory disease, by contacting a tissue having inflammation with an agent that modulates the expression of gene(s) from Gs in the tissue. M1 is useful for detecting GCA; M2 is useful for modulating GA; M3 is useful for screening an agent capable of modulating GCA preferably in an inflammation in a tissue; M4 is useful for detecting an inflammation (especially chronic) in a tissue, an allergic response in a subject, exposure of a subject to a pathogen or sterile inflammatory disease (e.g. psoriasis, rheumatoid arthritis, glomerulonephritis, asthma, thrombosis, cardiac reperfusion injury, renal reperfusion injury, ARDS, adult respiratory distress syndrome, inflammatory bowel disease, Crohn's disease, ulcerative colitis, periodontal disease; also bacterial infection, viral infection, parasitic infection, protozoal infection, fungal infection and M5 is useful for treating one of the above conditions. The present sequence represents a gene differentially expressed in granulocytes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 2000 BP; 367 A; 659 C; 602 G; 372 T; 0 U; 0 Other;

Query Match	89.5%;	Score 1951;	DB 6;	Length 2000;
Best Local Similarity	99.8%;	Pred. No. 0;		
Matches 1984;	Conservative	Mismatches 0;	Indels 3;	Gaps 3;

QY	196	ATTTCTCTGGGCTCCGGGGCGGAGAGCTGCATCCAGAGGAGCGCTCCAGGAGCG	255
Db	1	ATTTCTCTGGGCTCCGGGGCGGAGAGCTGCATCCAGAGGAGCGCTCCAGGAGCG	60
QY	256	GACCCGGAGTGTTTCAAGAGCCAGTGAACAGGACACAGGGGCCCAAGTCCACAGCCAT	315
Db	61	GACCCGGAGTGTTTCAAGAGCCAGTGAACAGGACACAGGGGCCCAAGTCCACAGCCAT	120
QY	316	GCAGACCTGCCCTCCCTGGCATTCCTTCCAGGCCCTTGGAGCCCTCTCTTT	375
Db	121	GCAGACCTGCCCTCCCTGGCATTCCTTCCAGGCCCTTGGAGCCCTCTCTTT	180
QY	376	TTTGCTGCTCTTGTAGTGTCTCAGATGAAGGCTGGGACAGCCCATCTGCACAGAGG	435
Db	181	TTTGCTGCTCTTGTAGTGTCTCAGATGAAGGCTGGGACAGCCCATCTGCACAGAGG	240
QY	436	GGTAGTCTCTGTGTCTTGGGGCGAGAACACCGTCAATGTCTTGCAACATCTCAAGCCTT	495
Db	241	GGTAGTCTCTGTGTCTTGGGGCGAGAACACCGTCAATGTCTTGCAACATCTCAAGCCTT	300
QY	496	CTCCCATGTCAATCAAGCTCGTCCCAAGGAGGAGCGCCATCTTCAATGAGGT	555
Db	301	CTCCCATGTCAATCAAGCTCGTCCCAAGGAGGAGCGCCATCTTCAATGAGGT	360
QY	556	GGCTCCAGGCTACTTCTCCCGGGACGGCTGGCAGCTCCAGGTTTCAGGGAGGCGGTGGCACA	615

Db 361 ||||| GGCCTCAGGCTACTTCTCCCGGACGCGCTGGAGCTCCAGGTTCAAGGGAGCGGTGGCACA 420
QY 616 GCTGTGTGATCAAAAGGCGCCGGGACTCCCATATGCTGGGCTGTACATGTGTGCACCTCTGTGGG 675
Db 421 GCTGTGTGATCAAAAGGCGCCGGGACTCCCATATGCTGGGCTGTACATGTGTGCACCTCTGTGGG 480
QY 676 ACACAGAGAAATAACAGACAAGTACAGCTGTGAGGTTTCAGGTGAGAAACCCAGTCCGC 735
Db 481 ACACAGAGAAATAACAGACAAGTACAGCTGTGAGGTTTCAGGTGAGAAACCCAGTCCGC 540
QY 736 CCCTGACACTGGGTTCTGGCTGTGCGAGCGGTGGTCACTGTCTTTCATCTTTGGT 795
Db 541 CCCTGACACTGGGTTCTGGCTGTGCGAGCGGTGGTCACTGTCTTTCATCTTTGGT 600
QY 796 CGCTCTGTGTCATGTTCCCTGGTACAGGTGCGGCTGTTCCAGCAACCGCGGAGAGAA 855
Db 601 CGCTCTGTGTCATGTTCCCTGGTACAGGTGCGGCTGTTCCAGCAACCGCGGAGAGNA 660
QY 856 GTTCTTCTCTAGAAACCCAGATGAAGTGTGAGCCCTTCAGAGCGGAGCCAGAGGG 915
Db 661 GTTCTTCTCTAGAAACCCAGATGAAGTGTGAGCCCTTCAGAGCGGAGCCAGAGGG 720
QY 916 CCGTGTGAGAGCTCGCTGAACTGTGGACCCAGACTCCGAGCCACCCCAAGCCGCT 975
Db 721 CCGTGTGAGAGCTCGCTGAACTGTGGACCCAGACTCCGAGCCACCCCAAGCCGCT 780
QY 976 GGCACTGTGTTCAAAACCTCACCACCTTGGAGCCCTGGAGCTGTGTTCCCTCCCAACCT 1035
Db 781 GGCACTGTGTTCAAAACCTCACCACCTTGGAGCCCTGGAGCTGTGTTCCCTCCCAACCT 839
QY 1036 TGTTCCTATATCCGAGACCATAGCCCTGCAAGGAGAGAGAGACACAGAGAGCA 1095
Db 840 TGTTCCTATATCCGAGACCATAGCCCTGCAAGGAGAGAGAGACACAGAGAGCA 899
QY 1096 GCGCTGAGTGCGACCTTGGGTGGCGGGCCCTGGGTCTCTGCTCCACCCGAGGGCACA 1155
Db 900 GCGCTGAGTGCGACCTTGGGTGGCGGGCCCTGGGTCTCTGCTCCACCCGAGGGCACA 959
QY 1156 GACACGGCTTGTGGCAGGCTGGGCTCTGTGTACCCACTCTCTGGTGGGTGCGAGAC 1215
Db 960 GACACGGCTTGTGGCAGGCTGGGCTCTGTGTACCCACTCTCTGGTGGGTGCGAGAC 1019
QY 1216 CTTTCCCTTCCACCCGCCAGGCTTCCAAAGCTCTGCTTCTCAGTTTCCAAATGGAAC 1275
Db 1020 CTTTCCCTTCCACCCGCCAGGCTTCCAAAGCTCTGCTTCTCAGTTTCCAAATGGAAC 1079
QY 1276 ACCTCACCCTCCGAGACCCGACTTACAGGACGATGCGCCCTCTGCTGCTCATCA 1335
Db 1080 ACCTCACCCTCCGAGACCCGACTTACAGGACGATGCGCCCTCTGCTGCTCATCA 1139
QY 1336 ACCCAGACCCGGACTCTCTTCTGCAACCCAGGCTGCTCGGCCCCAGGTTGGGT 1395
Db 1140 ACCCAGACCCGGACTCTCTTCTGCAACCCAGGCTGCTCGGCCCCAGGTTGGGT 1199
QY 1396 CCGCTCTCTCCACTCCAGGGCTCCGCGCCCAAGTGAAGGGGCCCTTGCAGAGCCTCAG 1455
Db 1200 CCGCTCTCTCCACTCCAGGGCTCCGCGCCCAAGTGAAGGGGCCCTTGCAGAGCCTCAG 1259
QY 1456 ACACACTGAGTTTACGGGCTGGGGGCGCTTGGCACAATCTGCTTGGCTATGAGCA 1515
Db 1260 ACACACTGAGTTTACGGGCTGGGGGCGCTTGGCACAATCTGCTTGGCTATGAGCA 1319
QY 1516 GCGTTTGGGGGCGCTTCCGCGGACGCCCGCGGGCGGAGTAGGTCTGGGGGCTTAGAG 1575
Db 1320 GCGTTTGGGGGCGCTTCCGCGGACGCCCGCGGGCGGAGTAGGTCTGGGGGCTTAGAG 1379
QY 1576 GCTGGGATGGCTCTGTGGCCCCACCGCCAGGGGGCAGCGCAGGCGGGCTGGAGCGGC 1635
Db 1380 GCTGGGATGGCTCTGTGGCCCCACCGCCAGGGGGCAGCGCAGGCGGGCTGGAGCGGC 1439
QY 1636 GCGGCGGCTCGGCTGGGGGTGAGTGGACGCTCTCGGGGCTGTTCGCGCATCC 1695

Db 1440 GCGCGCGCTCGGGCTCGGGGTGAGTGAGCGCTCCGGGGTGTGTCGCGCATCC 1499
QY 1696 TCAGTCCCTCGGCCACCCGGGGTGGTCCCTCGTGGCCACCGCACCTGCGCGAGCTCTT 1755
Db 1500 TCAGTCCCTCGGCCACCCGGGGTGGTCCCTCGTGGCCACCGCACCTGCGCGAGCTCTT 1559
QY 1756 TGGACCCAGATCTGTTCATGCTTTTGTCTGTCATCTGCGGGGGGCCCTTTGATGTCTT 1815
Db 1560 TGGACCCAGATCTGTTCATGCTTTTGTCTGTCATCTGCGGGGGGCCCTTTGATGTCTT 1619
QY 1816 CATCTGTATGGGTGGAATAATCACCGGGAATCCCTTCACTTCTTTGAAAAAGTTCCA 1875
Db 1620 CATCTGTATGGGTGGAATAATCACCGGGAATCCCTTCACTTCTTTGAAAAAGTTCCA 1679
QY 1876 TGACTCGAATATCTGAAATGAAGAAACAAACCGACTCACAAACCTCCAAAGTAGCTCCAA 1935
Db 1680 TGACTCGAATATCTGAAATGAAGAAACAAACCGACTCACAAACCTCCAAAGTAGCTCCAA 1739
QY 1936 ATGCAATTTTAAATGGAACAAATAATCTGAAAGAAACGCTTTAGTGGCTTTAAGCC 1995
Db 1740 ATGCAATTTTAAATGGAACAAATAATCTGAAAGAAACGCTTTAGTGGCTTTAAGCC 1799
QY 1996 CCATAACCTCCTTAAGCGTCTCGAGATGAAGACGGGGGGAG-CCCCAGCCAGGTGA 2054
Db 1800 CCATAACCTCCTTAAGCGTCTCGAGATGAAGACGGGGGGAGCCCCAGCCAGGTGA 1859
QY 2055 GACCCCGCAGGACCGCGCGCGCTCGAGCTGACCGAGGCTCGCACAGCGCGCGCTGAG 2114
Db 1860 GACCCCGCAGGACCGCGCGCGCTCGAGCTGACCGAGGCTCGCACAGCGCGCGCTGAG 1919
QY 2115 GGTGCGGCC-GAGCCAGGGTCCAAGAGGGGCGCTTGTGTCTCGGGTTAAATAGGTT 2173
Db 1920 GGTGCGGCCGAGCCAGGGTCCAAGAGGGGCGCTTGTGTCTCGGGTTAAATAGGTT 1979
QY 2174 CCGTCCG 2180
Db 1980 CCGTCCG 1986

RESULT 5

ADJ74845
ID ADJ74845 standard; DNA; 2000 BP.

XX AC ADJ74845;

XX DT 20-MAY-2004 (first entry)

XX DE Marker gene SEQ ID NO:97.

XX KW bronchial asthma; chronic obstructive pulmonary disease;

XX KW respiratory epithelial cell; interleukin-13; respiratory; antiasthmatic;

XX KW gene therapy; marker gene; gene; ds.

XX OS Homo sapiens.

XX PN EPI394274-A2.

XX PD 03-MAR-2004.

XX PF 04-AUG-2003; 2003EP-00254857.

XX PR 06-AUG-2002; 2002JP-00229312.

XX PR 20-MAR-2003; 2003JP-00077212.

XX PA (GENO-) GENOX RES INC.

XX PI Ohtani N, Sugita Y, Yamaya M, Kubo H, Nagai H, Izuhara K;

XX DR WPI; 2004-193155/19.

XX PT Testing for bronchial asthma or chronic obstructive pulmonary disease by
XX PT comparing the expression level of a marker gene in a biological sample
XX PT from a subject with the expression level of the gene in a sample from a

healthy subject.

Claim 1; SEQ ID NO 97; 241pp; English.

The present invention describes a method of testing for bronchial asthma or chronic obstructive pulmonary disease. The method comprises determining the expression level of a marker gene in a biological sample from a subject, comparing the expression level determined with the expression level of the marker gene in a biological sample from a healthy subject, and judging whether the subject has bronchial asthma or chronic obstructive pulmonary disease. The marker gene comprises: (a) a group of genes (S1) whose expression levels increase when respiratory epithelial cells are stimulated with interleukin-13; or (b) a group of genes (S2) whose expression levels decrease when respiratory epithelial cells are stimulated with interleukin-13. Also described: (1) a reagent (I) for testing for bronchial asthma or chronic obstructive pulmonary disease; (2) a kit for screening for a candidate compound for a therapeutic agent to treat bronchial asthma or chronic obstructive pulmonary disease; (3) an animal model for bronchial asthma or chronic obstructive pulmonary disease; (4) an inducer that induces bronchial asthma in a mouse; (5) a method for producing an animal model for bronchial asthma or chronic obstructive pulmonary disease; (6) a therapeutic agent for bronchial asthma or chronic obstructive pulmonary disease, comprising the compound, a marker gene or an antisense nucleic acid corresponding to a portion of the marker gene, a ribozyme, a polynucleotide that suppresses the expression of the gene through an RNAi effect or an antibody recognising a protein encoded by a marker gene; and (7) a DNA chip for testing for bronchial asthma or a chronic obstructive pulmonary disease, on which a probe has been immobilised to assay a marker gene. (I) has respiratory and antiasthmatic activities, and can be used in gene therapy. The method is useful for testing for or screening for a therapeutic agent for bronchial asthma or chronic obstructive pulmonary disease. The present sequence is used in the exemplification of the present invention.

Sequence 2000 BP; 367 A; 659 C; 602 G; 372 T; 0 U; 0 Other;

Query Match 89.58; Score 1951; DB 12; Length 2000;
Best Local Similarity 99.88; Pred. No. 0;
Matches 1984; Conservative 0; Mismatches 0; Indels 3; Gaps 3;

196 ATTTTCTGGGCTCCGGGCGGAGAGCTGCATCCAGAGAGCGGTCCAGAGCG 255
1 ATTTTCTGGGCTCCGGGCGGAGAGCTGCATCCAGAGAGCGGTCCAGAGCG 60

256 GACCCGGAGTGTTCAGAGCAGTGCACAGAGCAGGCGCCAGTCCACAGCCAT 315
61 GACCCGGAGTGTTCAGAGCAGTGCACAGAGCAGGCGCCAGTCCACAGCCAT 120

316 GCAGACCTGCCCTCCCTGGCATTCCTGGCCAGTTCCTCCAGGCGCTTGGGACCTCCTGTT 375
121 GCAGACCTGCCCTCCCTGGCATTCCTGGCCAGTTCCTCCAGGCGCTTGGGACCTCCTGTT 180

376 TTTGGCTCCTCTTGAAGTGTCTGAGATGAAGTGGGACAGCCCATCTGCAAGAGG 435
181 TTTGGCTCCTCTTGAAGTGTCTGAGATGAAGTGGGACAGCCCATCTGCAAGAGG 240

436 GGTAGTCTCTGTGTCTTGGGCGGAGACACCGTCTATGTCCTGCAACATCTCCAGCGCTT 495
241 GGTAGTCTCTGTGTCTTGGGCGGAGACACCGTCTATGTCCTGCAACATCTCCAGCGCTT 300

496 CTCCTCATGTCAACATCAAGCTCGGTGCCCGGAGAGAGCGCCATCTTCAATGAGGT 555
301 CTCCTCATGTCAACATCAAGCTCGGTGCCCGGAGAGAGCGCCATCTTCAATGAGGT 360

556 GGCTCCAGGCTACTTCTCCGGGACGGCTGGAGCTCAGGTTTCAAGGAGCGGTGGCACA 615
361 GGCTCCAGGCTACTTCTCCGGGACGGCTGGAGCTCAGGTTTCAAGGAGCGGTGGCACA 420

616 GCTGTGTATCAAGGCGCCCGGAGTCCCATGCTGGGCTGTACATGTGGCACCTCGTGGG 675
421 GCTGTGTATCAAGGCGCCCGGAGTCCCATGCTGGGCTGTACATGTGGCACCTCGTGGG 480

676 ACACCAGAGAAATAACAGAAAGTCAAGTCAAGTGGAGGTTTCAAGTGCAGAACCCAGTCCG 735

481 ACACCAGAGAAATAACAGAAAGTCAAGTGGAGGTTTCAGTGCAGAAACCCAGTCCGC 540

736 CCCTGACATCTGGGTTCTGGCCTGTGCAGCGGTGTCTACTGTCTTCTTCTTCTTGT 795

541 CCCTGACATCTGGGTTCTGGCCTGTGCAGCGGTGTCTACTGTCTTCTTCTTCTTGT 600

796 CGCTCTGGTCAATGTTCCGCTGTGTACAGTGTCCGCTGTTCCTCCAGCAACGCCGGGAGAGAA 855

601 CGCTCTGGTCAATGTTCCGCTGTGTACAGTGTCCGCTGTTCCTCCAGCAACGCCGGGAGAGAA 660

856 GTTCTTCTCTCTAGAAACCCAGATGAAGTTCGAGCCCTCAGAGCGGAGCCAGCAGGG 915

661 GTTCTTCTCTCTAGAAACCCAGATGAAGTTCGAGCCCTCAGAGCGGAGCCAGCAGGG 720

916 CTTGAGCAGAGCCTCCGCTGAACTGTGGACCCAGACTCCGAGCCCAACCCCAAGGCGCT 975

721 CTTGAGCAGAGCCTCCGCTGAACTGTGGACCCAGACTCCGAGCCCAACCCCAAGGCGCT 780

976 GGCATGCTGTCTCAAAACCCCTCAACCTTGGAGCCCTGGAGCTGTCTTCTTCTTCTTCTT 1035

781 GGCATGCTGTCTCAAAACCCCTCAACCTTGGAGCCCTGGAGCTGTCTTCTTCTTCTTCTT 839

1036 TGTTCATATGCGGAGAACCCATAGCCGCTGCAAGGAGAGAGAGAGAGAGAGAGAGAGAG 1095

840 TGTTCATATGCGGAGAACCCATAGCCGCTGCAAGGAGAGAGAGAGAGAGAGAGAGAGAG 899

1096 GCGCTGAGTGGGAGCTTGGGCTGGGCGGCTGGGCTCTCTGCTCCACCCCGGAGGAGCACA 1155

900 GCGCTGAGTGGGAGCTTGGGCTGGGCGGCTGGGCTCTCTGCTCCACCCCGGAGGAGCACA 959

1156 GACACCGGCTGTGTGGAGCGTGGGCTCTGTGTTCACCACTCTCTGGGCTGTCTGAGAG 1215

960 GACACCGGCTGTGTGGAGCGTGGGCTCTGTGTTCACCACTCTCTGGGCTGTCTGAGAG 1019

1216 CTTTCCCTTCAACCCCGGAGTCTTCAAGCTGTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1275

1020 CTTTCCCTTCAACCCCGGAGTCTTCAAGCTGTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1079

1276 ACCTCACCTCCGAGCAGCAGCCGACTTACAGAGAGCATGCCCCCTCTCTGCTTCTTCTT 1335

1080 ACCTCACCTCCGAGCAGCAGCCGACTTACAGAGAGCATGCCCCCTCTCTGCTTCTTCTT 1139

1336 ACCACAGACCGGAGCTCCCTTCTTCCACCCAGGCTGTCTCCGCGGCGGCTTGGGCT 1395

1140 ACCACAGACCGGAGCTCCCTTCTTCCACCCAGGCTGTCTCCGCGGCGGCTTGGGCT 1199

1396 CGGCTCTCTTCACTCCAGAGCTCCGCGCCCAAGTGGAGGCGGCGGCTTCCGAGAGCTCAG 1455

1200 CGGCTCTCTTCACTCCAGAGCTCCGCGCCCAAGTGGAGGCGGCGGCTTCCGAGAGCTCAG 1259

1456 ACACATGGAGTTCAGGCTTGGGCGGCTTGGGAGCATACCTGTCTTCTTCTTCTTCTTCTT 1515

1260 ACACATGGAGTTCAGGCTTGGGCGGCTTGGGAGCATACCTGTCTTCTTCTTCTTCTTCTT 1319

1516 GCGCTTGGGCGGCTTCCGCGGCTTCCGCGGCGGAGTAGGCTCTGGGCGGCTTACAG 1575

1320 GCGCTTGGGCGGCTTCCGCGGCTTCCGCGGCGGAGTAGGCTCTGGGCGGCTTACAG 1379

1576 GCTGGAGTGGCTCTCTTGGGCGGCTTGGGAGCATACCTGTCTTCTTCTTCTTCTTCTTCTT 1635

1380 GCTGGAGTGGCTCTCTTGGGCGGCTTGGGAGCATACCTGTCTTCTTCTTCTTCTTCTTCTT 1439

1636 GCGCGGCGGCTTCCGCGGCTTCCGCGGCTTCCGCGGCGGAGTAGGCTCTGGGCGGCTTACAG 1695

1440 GCGCGGCGGCTTCCGCGGCTTCCGCGGCTTCCGCGGCGGAGTAGGCTCTGGGCGGCTTACAG 1499

1696 TCAAGTCTCTTGGGCGGCTTCCGCGGCTTCCGCGGCGGAGTAGGCTCTGGGCGGCTTACAG 1755

1500 TCAAGTCTCTTGGGCGGCTTCCGCGGCTTCCGCGGCGGAGTAGGCTCTGGGCGGCTTACAG 1559

1756 TGAACCCAGATCTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1815

Db	721	CCTGAGCAGAGCCTCCGCTGAACCTGTGACCCACAGACTCCGAGCCACCCCAAGCCGCT	780
Qy	976	GGCACTGGTGTTCACAAACCTCACCACCTTTGGAGCCCTGGAGCTGTGTCTCCCCCAACCCCT	1035
Db	781	GGCACTGGTGTTCACAAACCTCACCACCTTTGGAGCCCTGGAGCTGTGTCTCCCCCAACCCCT	839
Qy	1036	TGTTTCCATATGCGCAGACCCATAGCCCTGCAAGCAGAGGACACAGGAGGCCA	1095
Db	840	TGTTTCCATATGCGCAGACCCATAGCCCTGCAAGCAGAGGACACAGGAGGCCA	899
Qy	1096	GGCCTGAGTCCGACCTTGGGTGGCGGGCTGCTCTGCTCCACCCCGAGGGGCACA	1155
Db	900	GGCCTGAGTCCGACCTTGGGTGGCGGGCTGCTCTGCTCCACCCCGAGGGGCACA	959
Qy	1156	GACACGGCTTCTTGGCAGGCTGGGCTCTGTGTACCCACTCTCGGTGGTGGTGCAGAC	1215
Db	960	GACACGGCTTCTTGGCAGGCTGGGCTCTGTGTACCCACTCTCGGTGGTGGTGCAGAC	1019
Qy	1216	CCTTCCCTCCACCCCGCAGGCTTCCAGCTCTGCTTCCAGTCTCTGCTCCAGTTCCTCAAAATGGAACC	1275
Db	1020	CCTTCCCTCCACCCCGCAGGCTTCCAGCTCTGCTTCCAGTCTCTGCTCCAGTTCCTCAAAATGGAACC	1079
Qy	1276	ACCTACCTCCGAGCACCAGCTTACACAGAGCATGCCCTCCCTCTGTGCCCTCATCAA	1335
Db	1080	ACCTACCTCCGAGCACCAGCTTACACAGAGCATGCCCTCCCTCTGTGCCCTCATCAA	1139
Qy	1336	ACCACAGACCCGGAATCCCTTTCTGCCACCCAGGCTGGTCCGGCCAGGTGTGGGT	1395
Db	1140	ACCACAGACCCGGAATCCCTTTCTGCCACCCAGGCTGGTCCGGCCAGGTGTGGGT	1199
Qy	1396	CCGCTCTCTCCACTCCAGGCTCCGCGCCCAAGTGAGGGGCCCTCGCGGAGCTCAG	1455
Db	1200	CCGCTCTCTCCACTCCAGGCTCCGCGCCCAAGTGAGGGGCCCTCGCGGAGCTCAG	1259
Qy	1456	ACACACTGGAGTTTCAGGGCTGGGGGGCCCTTGSCACATACCTGTCTCTGGCTATGAGCA	1515
Db	1260	ACACACTGGAGTTTCAGGGCTGGGGGGCCCTTGSCACATACCTGTCTCTGGCTATGAGCA	1319
Qy	1516	GGCTTTGGGGCCCTTCGCGGCGAGCCCGGGGGCCGAGGTAGGTCTGGGGCTTAGAG	1575
Db	1320	GGCTTTGGGGCCCTTCGCGGCGAGCCCGGGGGCCGAGGTAGGTCTGGGGCTTAGAG	1379
Qy	1576	GCTGGGATGCTCTCTGSCCCCAACCGCAGGGGCAAGCGAGGCGGGCTGGAGGCGGC	1635
Db	1380	GCTGGGATGCTCTCTGSCCCCAACCGCAGGGGCAAGCGAGGCGGGCTGGAGGCGGC	1439
Qy	1636	GGCGGCGGCTCGGGCTGGGGGTCAAGTGACGCTGCTCCGGGGCTGGTCGGGATCCC	1695
Db	1440	GGCGGCGGCTCGGGCTGGGGGTCAAGTGACGCTGCTCCGGGGCTGGTCGGGATCCC	1499
Qy	1696	TCAGTCCCTCGGCCACCCGGGGTCCGCTCCCTCGTCCACCGCAGCTGCGGAGCTCTT	1755
Db	1500	TCAGTCCCTCGGCCACCCGGGGTCCGCTCCCTCGTCCACCGCAGCTGCGGAGCTCTT	1559
Qy	1756	TGGACCCAGATCTCTTCATGCTTTTGTCTTCGTCACCTGCGCGGGGCCCTTTGATGTCT	1815
Db	1560	TGGACCCAGATCTCTTCATGCTTTTGTCTTCGTCACCTGCGCGGGGCCCTTTGATGTCT	1619
Qy	1816	CATCTGTATGGGTGGAAAAATCACCGGGAATCCCTTCAGTCTTTGAAAAAGTTCCA	1875
Db	1620	CATCTGTATGGGTGGAAAAATCACCGGGAATCCCTTCAGTCTTTGAAAAAGTTCCA	1679
Qy	1876	TGACTCGAATCTGAAATGAAGAAACAAACCGACTCACAAACCTCCAAAGTAGTCCAA	1935
Db	1680	TGACTCGAATCTGAAATGAAGAAACAAACCGACTCACAAACCTCCAAAGTAGTCCAA	1739
Qy	1936	ATGCAATTTTAAATGAAACAAAAATCTGAAAGAAACGCTTTTATGGCTTTAAGCC	1995
Db	1740	ATGCAATTTTAAATGAAACAAAAATCTGAAAGAAACGCTTTTATGGCTTTAAGCC	1799
Qy	1996	CAAAACCTCCCTAAGGCTCTCGAGTGAAGACGGGGGGAG-CCCAAGCCAGGTGGA	2054
Db	1800	CAAAACCTCCCTAAGGCTCTCGAGTGAAGACGGGGGGAGCCCCCAGCCAGGTGGA	1859

Qy	2055	GACCCCGCAGGACGCGCGCGCCCGGTGACCCGAGGCTTCGACAGCCGCGCCCTGTGAG	2114
Db	1860	GACCCCGCAGGACGCGCGCGCCCGGTGACCCGAGGCTTCGACAGCCGCGCCCTGTGAG	1919
Qy	2115	GGTCCGGGCC- GAGCCAGGGTCCAAAGGGGCGGCTTTGTCTCGGGTTAAATAAGGTT	2173
Db	1920	GGTCCGGGCCGAGCCAGGGTCCAAAGGGGCGGCTTTGTCTCGGGTTAAATAAGGTT	1979
Qy	2174	CCGTCCG 2180	
Db	1980	CCGTCCG 1986	

RESULT 7
ADP24663
ID ADP24663 standard; cDNA; 2000 BP.
XX
AC ADP24663;
XX
DT 18-NOV-2004 (first entry)
XX
DE PRO polypeptide encoding cDNA SEQ ID NO:1841.
XX
KW ss; gene; PRO; antinflammatory; antiarthritic; antirheumatic;
KW immunosuppressive; osteopathic; antidiabetic; dermatological;
KW antipsoriatic; antiallergic; antiasthmatic; hepatotropic; respiratory;
KW gene therapy; immune system.
XX
OS Unidentified.
XX
PN WO2004041170-A2.
XX
PD 21-MAY-2004.
XX
PF 30-OCT-2003; 2003WO-US034312.
XX
PR 01-NOV-2002; 2002US-0423394P.
XX
PA (GETH) GENENTECH INC.
XX
PI Clark H, Schoenfeld J, Van Lookeren M, Williams PM, Wood WJ;
PI Wu TD;
XX
WFI 2004-419628/39.
P-PSDB; ADP24664.
XX
PT New PRO polypeptides and polynucleotides, useful for treating e.g.
PT erythematous, rheumatoid arthritis, diabetes mellitus, immune-mediated
PT renal disease, or demyelinating diseases of the central or peripheral
PT nervous system.
XX
PS Claim 1; SEQ ID NO 1841; 2940pp; English.
XX
CC The invention relates to a novel isolated nucleic acid and the PRO
CC polypeptide encoded by it. A protein of the invention has
CC antinflammatory, antiarthritic, antirheumatic, immunosuppressive,
CC osteopathic, antidiabetic, dermatological, antipsoriatic, antiallergic,
CC antiasthmatic, hepatotropic, and respiratory activity. A polynucleotide
CC of the invention may have a use in gene therapy. The PRO polypeptide, its
CC agonist, antagonist, or antibody that specifically binds to the
CC polypeptide is useful for treating an immune related disorder such as
CC systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis,
CC juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis, an
CC idiopathic inflammatory myopathy, Sjogren's syndrome, autoimmune
CC vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune
CC thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal
CC disease, a demyelinating disease of the central or peripheral nervous
CC system, idiopathic demyelinating polyneuropathy, Guillain-Barre syndrome,
CC a chronic inflammatory demyelinating polyneuropathy, a hepatobiliary
CC disease, infectious or autoimmune chronic active hepatitis, primary
CC biliary cirrhosis, granulomatous hepatitis, sclerosing cholangitis,
CC inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's

CC	disease, an autoimmune or immune-mediated skin disease, a bullous skin
CC	disease, erythema multiforme, contact dermatitis, psoriasis, an allergic
CC	disease, asthma, allergic rhinitis, atopic dermatitis, food
CC	hypersensitivity, urticaria, an immunologic disease of the lung,
CC	eosinophilic pneumonia, idiopathic pulmonary fibrosis, hyperesensitivity
CC	pneumonitis, a transplantation associated disease, graft rejection or
CC	graft-versus-host disease. The present sequence encodes a PRO protein of
CC	the invention.
XX	
SQ	Sequence 2000 BP; 367 A; 659 C; 602 G; 372 T; 0 U; 0 Other;
	Query Match 89.5%; Score 1951; DB 13; Length 2000;
	Best Local Similarity 99.8%; Pred. No. 0;
	Matches 1984; Conservative 0; Mismatches 0; Indels 3; Gaps 3;
QY	196 ATTTCCTCGGGCTCCGGGCGGAGAGAGCTGCATCCAGAGGAGCGGTCCAGAGCG 255
DB	1 ATTTCCTCGGGCTCCGGGCGGAGAGAGCTGCATCCAGAGGAGCGGTCCAGAGCG 60
QY	256 GACCCGGAGTGTTCAGAGCAGTGAACAAGACAGAGGCGCCAAAGTCCCAACAGCCAT 315
DB	61 GACCCGGAGTGTTCAGAGCAGTGAACAAGACAGAGGCGCCAAAGTCCCAACAGCCAT 120
QY	316 GCAGACCTGCCCTCGGATTCCTCGGCCACGTTTCCAGAGCGCTTGGAGCCCTCTGTT 375
DB	121 GCAGACCTGCCCTCGGATTCCTCGGCCACGTTTCCAGAGCGCTTGGAGCCCTCTGTT 180
QY	376 TTGTGCTGCTCTTGAAGTGTCCAGAAAGAGCTGGAGACGCCCATCTGCACAGAGGG 435
DB	181 TTGTGCTGCTCTTGAAGTGTCCAGAAAGAGCTGGAGACGCCCATCTGCACAGAGGG 240
QY	436 GGTAGTCTCTGTGTGTTGGGGCGAAGAACCGTCAATGCTCTGCAACATCTCAAGCCCTT 495
DB	241 GGTAGTCTCTGTGTGTTGGGGCGAAGAACCGTCAATGCTCTGCAACATCTCAAGCCCTT 300
QY	496 CTCCTCATGTCAACATCAAGTCTGCTGCCACCGGAGGAGCGGCATCTCAATGAGGT 555
DB	301 CTCCTCATGTCAACATCAAGTCTGCTGCCACCGGAGGAGCGGCATCTCAATGAGGT 360
QY	556 GGCTCCAGGCTACTTCTCCGGGAGCGGTGGCAGCTTCCAGGTTTCAAGGAGGCGTGCCACA 615
DB	361 GGCTCCAGGCTACTTCTCCGGGAGCGGTGGCAGCTTCCAGGTTTCAAGGAGGCGTGCCACA 420
QY	616 GTGTGTATCAAGGCGCCCGGACTCCCATGCTGGGCTGTACATGTGGCACCTCGTGGG 675
DB	421 GTGTGTATCAAGGCGCCCGGACTCCCATGCTGGGCTGTACATGTGGCACCTCGTGGG 480
QY	676 ACACAGAGAAATACAGACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 735
DB	481 ACACAGAGAAATACAGACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 540
QY	736 CCTTGACACTGGGTTCTGGCTGTGCGAGCGTGGTCACTGTGTCTTTCATCTCTTGTGT 795
DB	541 CCTTGACACTGGGTTCTGGCTGTGCGAGCGTGGTCACTGTGTCTTTCATCTCTTGTGT 600
QY	796 CGCTCTGTGTCATGTTCGCTGTGACAGTGCCTGTTCCAGCAACCGCGGAGAGAA 855
DB	601 CGCTCTGTGTCATGTTCGCTGTGACAGTGCCTGTTCCAGCAACCGCGGAGAGAA 660
QY	856 GTTCTTCTCTAGAACCCAGATGAGTTCGAGCCCTCAGAGCGGAGGCCACAGGG 915
DB	661 GTTCTTCTCTAGAACCCAGATGAGTTCGAGCCCTCAGAGCGGAGGCCACAGGG 720
QY	916 CTTGAGCAGAGCTCTCGCTGAATGTGGACCCAGACTCCGAGCCACCCCAAGCCGCT 975
DB	721 CTTGAGCAGAGCTCTCGCTGAATGTGGACCCAGACTCCGAGCCACCCCAAGCCGCT 780
QY	976 GGCACCTGTGTTCAAAACCTTCAACCTTGGAGCCCTGGAGCTGCTGTCCCGCCCAACCT 1035
DB	781 GGCACCTGTGTTCAAAACCTTCAACCTTGGAGCCCTGGAGCTGCTGT - CCCCCCAACCT 839
QY	1036 TGTTCATATGCCGAGACCCATAGCCGCTTGAGGCGAGAGGACACAGGAGGCA 1095

DB	840 TGTTCCTCATATGCCCGCAGACCCATAGCCGCTTGAAAGCAGAGAGGACACAGGAGGCA 899
QY	1096 GGCCTGAGTGCAGACTTGGGTGGCGGGCTCTGGGTCTCTCTCCACCCGAGGAGGCA 1155
DB	900 GGCCTGAGTGCAGACTTGGGTGGCGGGCTCTGGGTCTCTCTCCACCCGAGGAGGCA 959
QY	1156 GACACCGGCTGTCTTGGCAGGCTGGGCTCTGTGTACCCACTCTCTGGGTGGGTGAGAC 1215
DB	960 GACACCGGCTGTCTTGGCAGGCTGGGCTCTGTGTACCCACTCTCTGGGTGGGTGAGAC 1019
QY	1216 CTTTCCTCCCTCCACCCCGAGGTCTTCCAAAGCTCTGCTTCTCTCAGTTTCCAAATGGAAC 1275
DB	1020 CTTTCCTCCCTCCACCCCGAGGTCTTCCAAAGCTCTGCTTCTCTCAGTTTCCAAATGGAAC 1079
QY	1276 ACCTCACCTCCGACACCCGACTTACAGAGCAGCATGCCCTCTCTCTGCCCTCATCAA 1335
DB	1080 ACCTCACCTCCGACACCCGACTTACAGAGCAGCATGCCCTCTCTCTGCCCTCATCAA 1139
QY	1336 ACCCAGACCCGAGCTCTCTTCTGTGCCACCCAGGCTGTTCGGGCCCGCAGGTGTGGGT 1395
DB	1140 ACCCAGACCCGAGCTCTCTTCTGTGCCACCCAGGCTGTTCGGGCCCGCAGGTGTGGGT 1199
QY	1396 CGGCTCTCTCACTCCAGGCTCCGCGCCCAAGTGAGGGGCGCCCTGCGGAGCCTCAG 1455
DB	1200 CGGCTCTCTCACTCCAGGCTCCGCGCCCAAGTGAGGGGCGCCCTGCGGAGCCTCAG 1259
QY	1456 ACACACTGGAGTTCAGGGCTGGGGGGCTTTGGGCACATACCTGTCTCTTGGCTATGAGCA 1515
DB	1260 ACACACTGGAGTTCAGGGCTGGGGGGCTTTGGGCACATACCTGTCTCTTGGCTATGAGCA 1319
QY	1516 GCTTTGGGGGCTTTCGCGGAGCGCCCGGGGGCGAGGTAGGTCTTGGGGCTTTAGAG 1575
DB	1320 GCTTTGGGGGCTTTCGCGGAGCGCCCGGGGGCGAGGTAGGTCTTGGGGCTTTAGAG 1379
QY	1576 GCTGGATGGCTCTCGGCCCCACCGCAGGGGGCAGCGAGCGGGCTGGGAGCGCGC 1635
DB	1380 GCTGGATGGCTCTCGGCCCCACCGCAGGGGGCAGCGAGCGGGCTGGGAGCGCGC 1439
QY	1636 GCGCGGCTCGGGCTGGGGGGCTCAGGTGGAGCTGCTCTCGGGGCTGTCTCGCGATCCC 1695
DB	1440 GCGCGGCTCGGGCTGGGGGGCTCAGGTGGAGCTGCTCTCGGGGCTGTCTCGCGATCCC 1499
QY	1696 TCAGTCCCTCGGCGCACCCGGGGCTCGCTCTGTGTGCCACCGCACCTGCGGAGCTCTT 1755
DB	1500 TCAGTCCCTCGGCGCACCCGGGGCTCGCTCTGTGTGCCACCGCACCTGCGGAGCTCTT 1559
QY	1756 TGGACCCAGATCTGTTCATGCTTTTGTCTTGTCTACTGGGGGGGCGCTTTGTGTCTT 1815
DB	1560 TGGACCCAGATCTGTTCATGCTTTTGTCTTGTCTACTGGGGGGGCGCTTTGTGTCTT 1619
QY	1816 CATCTGTATGGGTGGAAAAATCACCGGGAATCCCTCTTCAAGTTCTTTGAAAAAGTTCCA 1875
DB	1620 CATCTGTATGGGTGGAAAAATCACCGGGAATCCCTCTTCAAGTTCTTTGAAAAAGTTCCA 1679
QY	1876 TGACTCGAATATCTGAAATGAAAGAAAAAACAACCGACTCAAAAACCTCAAGTAGTCCAA 1935
DB	1680 TGACTCGAATATCTGAAATGAAAGAAAAAACAACCGACTCAAAAACCTCAAGTAGTCCAA 1739
QY	1936 ATGCAATTTTAAATGAAAGAAAAAATAATCTGAAGAAACGCTTTTGTGTCTTTAAGCC 1995
DB	1740 ATGCAATTTTAAATGAAAGAAAAAATAATCTGAAGAAACGCTTTTGTGTCTTTAAGCC 1799
QY	1996 CCAAAACCTCTCCTAAGGCTCTCTGAGATGAAGAGCGGGGGGAG - CCGCAGCAGGTGA 2054
DB	1800 CCAAAACCTCTCCTAAGGCTCTCTGAGATGAAGAGCGGGGGGAGCCCGCAGCAGGTGA 1859
QY	2055 GACCCCGAGAGCGCGCGCGCGGTGACCGAGGCTTCGACAGCGCGCGCGCTCTGAG 2114
DB	1860 GACCCCGAGAGCGCGCGCGCGGTGACCGAGGCTTCGACAGCGCGCGCGCTCTGAG 1919
QY	2115 GGTGGGCG - GAGCCAGGCTCCAGAGGGGCGGTTTGTGTCTCGGTTTAAATAGGTT 2173
DB	1920 GGTGGGCGGAGCGAGGCTCCAGAGGGGCGGTTTGTGTCTCGGTTTAAATAGGTT 1979

1096 GGCCTGAGTGCAGCCTTGGGTGGCGGGCCCTGGGTCTCTCTCCACCCCGAGGCGACA 1155
Db |||||||
895 GGCCTGAGTGCAGCCTTGGGTGGCGGGCATGGGTCTCTCTCCACCCCGAGGCGACA-CA 953
Qy |||||||
1156 GACACCGGCTTGGTGGCAGGCTGGCGCTCTGTGTACCCACTCTCTGGGTGGGTGCAGAC 1215
Db |||||||
954 GACACCGGCTTGGTGGCAGGCTGGCGCTCTGTGTACCCACTCTCTGGGTGGGTGCAGAC 1013
Qy |||||||
1216 CTTTCCCTCCACCCCGAGGCTTTCAGAGCTCTGTCTCTCAGTTTCCAAAATGGAACC 1275
Db |||||||
1014 CTTTCCCTCCACCCCGAGGCTTTCAGAGCTCTGTCTCTCAGTTTCCAAAATGGAACC 1073
Qy |||||||
1276 ACTCATCTCCGAGACACCCGACTTATCCAGAGCGCATGGCCCTCTCTCTGCTCATCAA 1335
Db |||||||
1074 AACTCACCTTCCGAGCAACCGAATACCAAGAACGATGCCCTCTCTCTGCTCATCAA 1133
Qy |||||||
1336 ACCACAGACCGGCTCTCTTCTTCCACCCAGGCTGGTCCGGCCCGCAGGTGGGT 1395
Db |||||||
1134 ACCACAGACCGGCTCTCTTCTTCCACCCAGGCTGGTCCGGCCCGCAGGTGT-GGGT 1192
Qy |||||||
1396 CCGCTCTCTCCACTCCAGGCTCCGCGCCCAAGTGAGGGGCCCTGCGGAGCTCAG 1455
Db |||||||
1193 CCGCTCTCTCCACTCCAGGCTCCGCG-CAAGTGAGGGGCCCTGCGGAGCTCAG 1251
Qy |||||||
1456 ACACACTGGAGTTCAGGGCTG-GGGGGCTTTGGCACATACCTGTCTCTTGGCTATGAGC 1514
Db |||||||
1252 ACACACTCCAGTTCCAGGCTGTGGGGGCTGCGCCACATACCTGTCTCTTGGCTATGAGC 1311
Qy |||||||
1515 AGGCTTTGGGGCCCTTCCGGCGACGCCCGGGGCGGAGGTAGGCTCTGGGGCTTAGA 1574
Db |||||||
1312 AGGCTTTGGGGCCCTTCCGGCGAG-CCCGGGGGCCGAGGTAGGTC-GGGGGCTTAGA 1369
Qy |||||||
1575 GCGTGGGATGGCTCTCTGGGCCCCACCGCCAGGGGCAAGCGAGGCGGGCTGGAGGGCG 1634
Db |||||||
1370 GCGTGGGATGGCTCTCTGGGCCCCACCGCCAGGGGCG-AGCGAGGCGGGCTGGAGGGCG 1428
Qy |||||||
1635 CGCGCGCGCTCGGGCTGGGGGTGAGGTGAGCGTG-CCTCGGGGCTGTGCGCGATC 1693
Db |||||||
1429 CGCGCGCGCTCGGGCTGGGGGTGAGGTGAGCGCGCGCCCTCGGGGCTGGAGCGCATC 1488
Qy |||||||
1694 CCTCAGTCCCTGGCCACCGGGGTGCTCTCTGTGCCACCGCACTGCGGAGCTC 1753
Db |||||||
1489 CCTCAGTCCCTGGCCACCGGGGTGCTCTCTGTGCCACCGCACTGCGGAGCTC 1548
Qy |||||||
1754 TTGACACCGAGATCTGTTTCATGCTTTTGTCTTCTGTCACCTGCGGGCGGGCCCTTTCATGTC 1813
Db |||||||
1549 TTGACACCGAGATCTGTTTCATGCTTTTGTCTTCTGTCACCTGCGGGCGGGCCCTTTCATGTC 1608
Qy |||||||
1814 TTCATCTGTATGGGTGGAAAAATCACCGGGAATCCCTTCAGTTCTTTGAAAAAGTTC 1873
Db |||||||
1609 TTCATCTGTATGGGTGGAAAAATCACCGGGAATCCCTTCAGTTCTTTGAAAAAGTTC 1668
Qy |||||||
1874 CATGACTCGAATATCTGAATGAAGAAACAAACCGACTCACAACCTCCAAAGTAGCTCC 1933
Db |||||||
1669 CATGACTCGAATATCTGAATGAAGAAACAAACCGACTCACAACCTCCAAAGTAGCTCC 1728
Qy |||||||
1934 AAATGCAATTTTAAATGGAACAAATACTGAAAGAAACGCTCTTAGTGGCTTTAAG 1993
Db |||||||
1729 AAATGCAATTTTAAATGGAACAAATACTGAAAGAAACGCTCTTAGTGGCTTTAAG 1788
Qy |||||||
1994 CCCCCAACCGTCCCTAAGCGCTCTCGAGATGAAGACGGGGGGAGCCCCCAGCGAGTGG 2053
Db |||||||
1789 CCCCCAACCGTCCCTAAGCGCTCTCGAGATGAAGACGGGGGGAG--CCAGCCAGGTGG 1846
Qy |||||||
2054 AGACCCCGAGACGGCGGGCGCGCTGAGCGAGGCTCTGCAAGCGCGCGCGCTTGA 2113
Db |||||||
1847 AGACCCCGAG--CGCGGGCGCGCGCTGAGCGAGGCTCTGCAAGCGCGCGCGCTTGA 1904
Qy |||||||
2114 GGGTGGGCGGACGAGGCTCAAGAGGGGCGGCTTGTGCTCGGTTAAATAGGTT 2173
Db |||||||
1905 GGGTGGGCGGACGAGGCTCAAGAGGGGCGGCTTGTGCTCGGTTAAATAGGTT 1964

Qy 2174 CCGTCCG 2180
Db |||||||
1965 CCGTCCG 1971

RESULT 9

ADQ85525
ID ADQ85525 standard; cDNA; 1985 BP.
XX AC ADQ85525;
XX DT 07-OCT-2004 (first entry)
XX Human tumour-associated antigenic target (TAT) cDNA sequence #2339.
DE human; tumour-associated antigenic target; TAT; cytostatic; gene therapy;
KW cancer; cell proliferative disorder; gene; ss.
XX Homo sapiens.
XX WO2004060270-A2.
XX 22-JUL-2004.
XX 15-OCT-2003; 2003WO-US029126.
XX 18-OCT-2002; 2002US-0418988P.
XX (GETH) GENENTECH INC.
XX (WUTD/) WU T D.
XX (ZHOU/) ZHOU Y.
XX Wu TD, Zhou Y;
XX WPI; 2004-534300/51.
XX New nucleic acid molecule and encoded polypeptide, for diagnosing,
preventing or treating cell proliferative disorders such as cancer.
XX Claim 1; SEQ ID NO 2339; 5504pp; English.
CC The present invention describes an isolated tumour-associated antigenic
target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide
sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of
CC (a); (c) the complement of (a) or (b); (d) a sequence that has 80%
CC sequence identity to (a)-(c); or (e) a sequence that hybridises to (a) -
CC (c). Also described: (1) an expression vector comprising the above
CC nucleic acid; (2) a host cell comprising the above expression vector; (3)
CC a process for producing a polypeptide; (4) an isolated polypeptide
CC comprising: (a) an amino acid sequence encoded by any of the above
CC nucleotide sequences; (b) an amino acid sequence encoded by the full-
CC length coding region of the above nucleotide sequences; or (c) a sequence
CC having at least 80% identical to (a) or (b); (5) a chimeric polypeptide
CC comprising the above polypeptide fused to a heterologous polypeptide; (6)
CC an isolated antibody that binds to the above polypeptide; (7) a process
CC for producing the antibody; (8) an isolated oligopeptide that binds to
CC the above polypeptide; (9) a tumour-associated antigenic target (TAT)
CC binding organic molecule that binds to the above polypeptide; (10) a
CC composition of matter comprising the above (chimeric) polypeptide,
CC antibody, oligopeptide or TAT binding organic molecule, in combination
CC with a carrier; (11) an article of manufacture comprising a container and
CC the composition of matter contained within the container; (12) methods of
CC inhibiting the growth of a cell that expresses the above protein, where
CC the growth of the cell is at least in part dependent upon a growth
CC potentiating effect of the above protein; (13) a method of
CC therapeutically treating a mammal having a cancerous tumour comprising
CC cells that express the above protein; (14) a method of determining the
CC presence of a protein in a sample suspected of containing the protein
CC described above; (15) methods of diagnosing the presence of a tumour in a
CC mammal; (16) a method for treating or preventing a cell proliferative
CC disorder associated with increased expression or activity of the above
CC protein; and (17) a method of binding an antibody, oligopeptide or
CC organic molecule to a cell that expresses the protein described above.

The TAT sequences have cytostatic activities, and can be used in gene therapy. The composition and methods are useful for diagnosing, preventing or treating cancer. The composition is also used for preparing a medicament for the therapeutic treatment or diagnostic detection of a cell proliferative disorder or cancer. The present sequence represents a human TAT cDNA sequence from the present invention.

XX
SQ
Semience 1985 BP: 372 A: 649 C: 594 G: 370 T: 0 U: 0 Other;

Query Match	80.3%;	Score 1751;	DB 13;	Length 1985;
Best Local Similarity	97.8%;	Pred. No. 0;		
Marches 1944: Conservative	0;	Mismatches 25;	Indels 18;	Gaps 16;

QY	196	ATTTTCTCTGGGCTCGGGCGCGGAGAGCTGCATCCACAGAGAGCGGTCCAGGAGCG	255
DB	1	ATTTTCTCTGGGCTCGGGCGCGGAGAGAGCTGCATCCAGAGGAGCGGTCCAGGAGCG	60
QY	256	GACCCGGGAGTGTTCAGAGCCAGTGCACAGGACACAGGGGCCAAGTCCACACAGCAT	315
DB	61	GACCCGGGAGTGTTCAGAGCCAGTGCACAGGAGCA--GGGCCCAAGTCCACACAGCAT	119
QY	316	GCAGACTGCCCTCGGCATTCCCTGGCCACAGCTTTCCAGGCGCTTGAGGACCTCTCTGTT	375
DB	120	GCAGACTGCCCTCGGCATTCCCTGGCCACAGCTTTCCAGGCGCTTGAGGACCTCTCTGTT	179
QY	376	TTTGGCTGCCTCTTGAGTGTCTCAGAAATGAAGCGCTGGGACAGCCCATCTCTGCACAGAGG	435
DB	180	TTTGGCTGCCTCTTGAGTGTCTCAGAAATGAAGCGCTGGGACAGCCCATCTCTGCACAGAGG	239
QY	436	GGTAGTCTCTGTGTCTTGGGGGAGAAACACCGTCACTGTCTTCCACACTCTCCACGCCCTT	495
DB	240	GGTAGTCTCTGTGTCTTGGGGGAGAAACACCGTCACTGTCTTCCACACTCTCCACGCCCTT	299
QY	496	CTCCATGTCAACATCAAGCTGTGTGCCACAGCGGTCACTGTCTTCCACACTCTCCACGCCCTT	555
DB	300	CTCCATGTCAACATCAAGCTGTGTGCCACAGCGGTCACTGTCTTCCACACTCTCCACGCCCTT	359
QY	556	GGCTCCAGGCTACTTCTCCCGGAGCGGTGGAGCTCCAGGTTCCAGGAGGGGTGGACACA	615
DB	360	GGCTCCAGGCTACTTCTCCCGGAGCGGTGGAGCTCCAGGTTCCAGGAGGGGTGGACACA	419
QY	616	GCTGGTGAACAAAGGCGCGGGAGCTCCCATGCTGGGCTGTACATGTGGCACTCTCGTGGG	675
DB	420	GCTGGTGAACAAAGGCGCGGGAGCTCCCATGCTGGGCTGTACATGTGGCACTCTCGTGGG	479
QY	676	ACACAGAGAAATAACAGACAAGTCACTGTGGAGGTTTCAGGTGCGAGAACCCACGATCCGC	735
DB	480	ACACAGAGAAATAACAGACAAGTCACTGTGGAGGTTTCAGGTGCGAGAACCCACGATCCGC	539
QY	736	CCCTGACACTGGTTCCTGGCTGTGCCAGCGGTGGTCACTGCTGTCTTCATCTCTTGGT	795
DB	540	CCCCGACACTGGGTT--TGCGCTGTGCCAGCGGTGGTCACTGCTGTCTTCATCTCTTGGT	598
QY	796	CGCTCTGTGTCTGTTCGCTCGGTACAGTGCCTGTTCAGCAACCGCGGAGAGAA	855
DB	599	CGCTCTGTGTCTGTTCGCTCGGTACAGTGCCTGTTCAGCAACCGCGGAGAGAA	657
QY	856	GTTCTTCTCTCTAGAAACCCAGATGAAGTTCGACGCTTCAGAGCGGAGCCACGAGG	915
DB	658	GTTCTTCTCTCTAGAAACCCAGATGAAGTTCGACGCTTCAGAGCGGAGCCACGAGG	716
QY	916	CTTGAGCAGAGCTCGGTGAACCTGTGGACCCACAGACTCCGAGGCCACCCCAAGGCGCT	975
DB	717	CTTGAGCAGCA--CCTCCGCTGAACTGTGACCCAGACTCCGAGGCCACCCCAAGGCGCT	775
QY	976	GGCACTGTGTGTTCAAAACCCCTACCACTTTGGAGCCCTGGAGTGTCTCCGCCCAACCT	1035
DB	776	GGCACTGTGTGTTCAAAACCCCTACCACTTTGGAGCCCTGGAGTGTCTGT--CCGCCCAACCT	834
QY	1036	TGTTTCCATATGCCCGAGACCATAGCGCTTGCAGGACGAGAGGACACAGGAGAGCCA	1095
DB	835	TGTTTCCATATGCCCGAGACCATAGCGCTTGCAGGACGAGAGGACACAGGAGAGCCA	894

Db	1965	CGTCCG 1971	
XX	RESULT 10		
XX	AAC59772		
XX	ID AAC59772 standard; cDNA; 1464 BP.		
XX	AC AAC59772;		
XX	26-JAN-2001 (first entry)		
XX	DE Human secreted protein gene 35 SEQ ID NO:45.		
XX	KW Human; secreted protein; diagnosis; immunosuppressive; antiarthritic;		
XX	KW antirheumatic; antiproliferative; cytostatic; cardiant; vasotropic;		
XX	KW cerebroprotective; neurotropic; neuroprotective; antibacterial; virucide;		
XX	KW fungicide; ophthalmological; gene therapy; autoimmune disease; neoplasm;		
XX	KW hyperproliferative disorder; cancer; cardiovascular disorder; infection;		
XX	KW cerebrovascular disorder; angiogenesis; nervous system disorder;		
XX	KW ocular disorder; wound healing; skin aging; food additive; preservative;		
XX	ss.		
XX	OS Homo sapiens.		
XX	XX WO200056751-A1.		
XX	XX 28-SEP-2000.		
XX	XX 09-MAR-2000; 2000WO-US006013.		
XX	XX 19-MAR-1999; 99US-0125360P.		
XX	XX 11-JUN-1999; 99US-0138626P.		
XX	XX 03-DEC-1999; 99US-0168662P.		
XX	XX (HUMA-) HUMAN GENOME SCI INC.		
XX	XX Rosen CA, Ruben SM, Komatsoulis G;		
XX	XX WPI; 2000-579482/54.		
XX	XX P-PSDB; AAB34611.		
XX	XX Isolated nucleic acid molecule encoding a human secreted protein is used		
XX	XX in preventing, treating or ameliorating a medical condition.		
XX	XX Claim 1; Page 362; 419pp; English.		
XX	XX The polynucleotide sequences given in AAC59738 to AAC59787 encode the		
XX	XX human secreted proteins given in AAB34577 to AAB34626. AAB34627 to		
XX	XX AAB34686 represent human secreted polypeptide sequences and proteins		
XX	XX homologous to them, which are given in the exemplification of the present		
XX	XX invention. Human secreted proteins have activities based on the tissues		
XX	XX and cells the genes are expressed in. Example of activities include:		
XX	XX antiarthritic; immunosuppressive; antirheumatic; antiproliferative;		
XX	XX cytostatic; cardiant; vasotropic; cerebroprotective; neurotropic;		
XX	XX neuroprotective; antibacterial; virucide; fungicide; and		
XX	XX ophthalmological. The polynucleotides and proteins can be used to		
XX	XX prevent, treat or ameliorate a medical condition in e.g. humans, mice,		
XX	XX rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used		
XX	XX in diagnosing a pathological condition or susceptibility to a		
XX	XX pathological condition. Disorders which are diagnosed or treated include		
XX	XX autoimmune diseases, hyperproliferative disorders e.g. neoplasms and		
XX	XX cancers of the breast or liver, cardiovascular disorders, cerebrovascular		
XX	XX disorders, angiogenesis, nervous system disorders, infections caused by		
XX	XX bacteria, viruses and fungi and ocular disorders. The proteins can also		
XX	XX be used to aid wound healing and epithelial cell proliferation, to		
XX	XX prevent skin aging due to sunburn, to maintain organs before		
XX	XX transplantation, for supporting cell culture of primary tissues, to		
XX	XX regenerate tissues and in chemotaxis. The proteins can also be used as a		
XX	XX food additive or preservative to increase or decrease storage		
XX	XX capabilities. AAC59729 to AAC59737 and AAB34576 represent sequences used		
XX	XX in the exemplification of the present invention		

SQ	Sequence 1464 BP; 309 A; 483 C; 408 G; 264 T; 0 U; 0 Other;
Query Match	43.1%; Score 938.6; DB 3; Length 1464;
Best Local Similarity	95.7%; Pred. No. 4.9e-205;
Matches 1040; Conservative	0; Mismatches 39; Indels 8; Gaps 7;
QY	881 AAGGTCGAGCCCTCAGAGCGGGAGCCAGCAGAGGCGCTGAGCAGAGCTCCCTGAACTG 940
DB	111 AGGCGCCCGGACTCAGAGCGGGAGCCAGCAGAGGCGCTGAGCAGAGCTCCGCTGAACTG 170
QY	941 TGGACCCGAGACTCCGAGGCCACCCAGGCGCTGGCACTGGTGTCAAAACCTCACCA 1000
DB	171 TGGACCCGAGACTCCGAGGCCACCCAGGCGCTGGCACTGGTGTCAAAACCTCACCA 230
QY	1001 CTTGAGAGCCCTGGAGCTGCTGCCCCCCCAACCCCTTTTCCATATGCCGAGAGCCATA 1060
DB	231 CTTGAGAGCCCTGGAGCTGCTGT-CCCCCAACCCCTTGGTTCATATGCCGAGAGCCATA 289
QY	1061 GCCCTCTGAAGGCAAGAGGACACAGAGAGCCAGCCCTGAGTCCGACCTTGGGTGGC 1120
DB	290 GCCCTCTGAAGGCAAGAGGACACAGAGAGCCAGCCCTGAGTCCGACCTTGGGTGGC 349
QY	1121 GGGGCTGGGTCTCTCGTCCACCCGAGGGGACAGACACCGGCTTGTGGCAGGCTGG 1180
DB	350 GGGGCTGGGTCTCTCGTCCACCCGAGGGGACAGACACCGGCTTGTGGCAGGCTGG 409
QY	1181 GCCTCTGTGTACCCCACTCTCGGGTGCCTGAGAGCCCTTCCCTCCACCCCGGCTCTT 1240
DB	410 GCCTCTGTGTACCCCACTCTCGGGTGCCTGAGAGCCCTTCCCTCCACCCCGGCTCTT 469
QY	1241 CCAAGCTCTGCTTCTCAGTTTCCAAATGGAACCACTCACTCCGAGACACCCGACTT 1300
DB	470 CCAAGCTCTGCTTCTCAGTTTCCAAATGGAACCACTCACTCCGAGACACCCGACTT 529
QY	1301 ACCAGGACGATGCCCTCTCCCTCTGCTCATAAACCCAGACCCGACTCCCTTCT 1360
DB	530 ACCAGGACGATGCCCTCTCCCTCTGCTCATAAACCCAGACCCGACTCCCTTCT 589
QY	1361 GCCACCCGAGGCTGGTCCGGGCCCGAGGTGGGGTCCGCTCTCTCCACTCCCGAGGCTCC 1420
DB	590 GCCACCCGAGGCTGGTCCGGGCCCGAGGTGGGGTCCGCTCTCTCCACTCCCGAGGCTCC 648
QY	1421 GCGCCCAAGTGAAGGGGGCCCTGCGGAGCCTCAGACACACTTGGAGTTGAGGCG-..TGGG 1478
DB	649 GCGCCCAAGTGAAGGGGGCCCTGCGGAGCCTCAGACACACTTGGAGTTGAGGCGTGGG 708
QY	1479 GGGGCTTGGCACATACCTGTCCCTTGGCTATGACAGAGCTTGGGGGCCCTTCCGCGG 1538
DB	709 GGGGCTTGGCACATACCTGTCCCTTGGCTATGACAGAGCTTGGGGGCCCTTCCGCGG 768
QY	1539 AGCCCGGGGGCCGAGGTAGGTCTGGGGGCTTAGAGGCTGGGATGGGTCTCTGGGCCCCAC 1598
DB	769 AGCCCGGGGGCCGAGGTAGGTCTGGGGGCTTAGAGGCTGGGATGGGTCTCTGGGCCCCAC 827
QY	1599 GCGCAGGGGGCAAGCGCAGGCGGGCTGGGAGGGCGGCGCGGCTCGGGCTGGGGGGT 1658
DB	828 GCGCAGGGGGG-AGCGCAGGCGGGCTGGGAGGGCGGCGGCGGCTCGGGCTGGGGGGT 886
QY	1659 CAGGTGAGAGCTG-CTTCGGGGCTGGTTCGGGCATCCCTCAGTCCCTCGGCCACCCGGGG 1717
DB	887 CAGGTGAGAGCTGCTCCCTCGGGGCTGGTTCGGGCATCCCTCAGTCCCTCGGCCACCCGGGG 946
QY	1718 GTGCGTCTCCTCGTGGCCACCGCAGCTCGGAGCCCTTTTGGACCCAGATCTGTTTCATGCT 1777
DB	947 GTGCGTCTCCTCGTGGCCACCGCAGCTCGGAGCCCTTTTGGACCCAGATCTGTTTCATGCT 1006
QY	1778 TTTGTCTTCTGTCAGTGGCGGGGCCCTTTGATGTCTTCATCTGTAT-GGGGTGGAAAAA 1836
DB	1007 TTTGTCTTCTGTCAGTGGCGGGGCCCTTTGATGTCTTCATCTGTATGGGGTGGAAAAA 1066
QY	1837 TCACCGGAATCCCGCTTTCAGTCTTTTGAAGGTTCCATGACTCGAATATCTGAAATGA 1896
DB	1067 TCACCGGAATCCCGCTTTCAGTCTTTTGAAGGTTCCATGACTCGAATATCTGAAATGA 1126

QY 1897 AGAAAAACCAACGACTCAACACCTCCCAAGTAGCTCCAAATGCAATTTTAAATGAAA 1956
Db 1127 AGAAAAACCAACGACTCAACAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 1186
QY 1957 ACAAAAA 1963
Db 1187 AAAAAAA 1193
RESULT 11
AAX82081
ID AAX82081 standard; DNA; 1064 BP.
AC AAX82081;
XX 20-SEP-1999 (first entry)
XX Human SIGP encoding DNA (clone ID 1747327).
XX Signal-peptide containing protein; SIGP; human; cancer; immune response;
XX adenocarcinoma; leukemia; lymphoma; melanoma; myeloma; sarcoma; AIDS;
KW Addison's disease; adult respiratory distress syndrome; allergy; anemia;
KW asthma; atherosclerosis; bronchitis; cholecystitis; Crohn's disease;
KW ulcerative colitis; atopic dermatitis; dermatomyositis; emphysema;
KW diabetes mellitus; atrophic gastritis; glomerulonephritis; gout; trauma;
KW Grave's disease; hyperesinophilia; irritable bowel syndrome; infection;
KW lupus erythematosus; multiple sclerosis; myasthenia gravis; inflammation;
KW osteoarthritis; osteoporosis; pancreatitis; polymyositis; scleroderma;
KW rheumatoid arthritis; Sjogren's syndrome; autoimmune thyroiditis; ss.
XX Homo sapiens.
OS
XX WO9933981-A2.
XX 08-JUL-1999.
XX 22-DEC-1998; 98WO-US027598.
XX 31-DEC-1997; 97US-00002485.
XX (INCY-) INCYTE PHARM INC.
XX Lal P. Hillman JL, Corley NC, Guegler KJ, Baughn MR, Sather SK;
PI Shah P;
XX WPI; 1999-430242/36.
XX P-PSDB; AAY21846.
XX Human signal-peptide containing protein coding sequences used to treat
XX cancer and immune responses.
XX Claim 9; Page 92-93; 99pp; English.
XX The invention provides human signal-peptide containing proteins (SIGP)
XX (AAY21841-855) and polynucleotides (AAX82076-90) encoding the proteins. A
XX host cell containing a vector comprising SIGP DNA can be used to produce
XX the SIGP protein. The SIGP protein can be used, in conjunction with a
XX pharmaceutical carrier to treat or prevent a cancer. An antagonist of the
XX SIGP protein can be used to treat or prevent a cancer or an immune
XX response. The cancers that can be treated or prevented include sarcomas,
XX adenocarcinomas, leukemia's, lymphomas, melanomas, teratocarcinomas,
XX myelomas and cancers of the adrenal gland, bladder, bone, bone marrow,
XX brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract,
XX heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis,
XX prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and
XX uterus. The immune responses that can be treated or prevented include,
XX AIDS, Addison's disease, adult respiratory distress syndrome, allergies,
XX anemia, asthma, atherosclerosis, bronchitis, cholecystitis, Crohn's
XX disease, ulcerative colitis, atopic dermatitis, dermatomyositis, diabetes
XX mellitus, emphysema, atrophic gastritis, glomerulonephritis, Grave's
XX disease, gout, hyperesinophilia, irritable bowel syndrome, lupus
XX erythematosus, multiple sclerosis, myasthenia gravis, myocardial or.

CC pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, and
CC polymyositis, rheumatoid arthritis, scleroderma, Sjogren's syndrome, and
CC autoimmune thyroiditis, complications of cancer, infection, and trauma
XX Sequence 1064 BP; 198 A; 349 C; 315 G; 202 T; 0 U; 0 Other;
SQ Query Match 41.9%; Score 912.8; DB 2; Length 1064;
Best Local Similarity 96.1%; Pred. No. 3.6e-199;
Matches 957; Conservative 0; Mismatches 37; Indels 2; Gaps 2;
QY 199 TTCTGGGGCTCCGGGGCGCGGAGAGCTGCATCCAGAGGAGCGGTCCAGGAGCGGAC 258
Db 1 TTCTGGGGCTCCGGGGCGCGGAGAGCTGCATCCAGAGGAGCGGTCCAGGAGCGGAC 60
QY 259 CCGGGAGTGTTCAGAGAGCCAGTGAACAAGAGCAGGGGGCCCAAGTCCCACAGCCATGCA 318
Db 61 CCGGGAGTGTTCAGAGAGCCAGTGAACAAGAGCAGGGGGCCCAAGTCCCACAGCCATGCA 120
QY 319 GACCTGCCCTCCCTGGCATTCCTCGGCCACGTTTCCAGGCCCTTTGGGACCCCTCTGTTTT 378
Db 121 GACCTGCCCTCCCTGGCATTCCTCGGCCACGTTTCCAGGCCCTTTGGGACCCCTCTGTTTT 180
QY 379 GGCTGCCCTCCCTGGCATTCCTCGGCCACGTTTCCAGGCCCTTTGGGACCCCTCTGTTTT 438
Db 181 GGCTGCCCTCCCTGGCATTCCTCGGCCACGTTTCCAGGCCCTTTGGGACCCCTCTGTTTT 240
QY 439 AGTCTCTGTGTCTTGGGGCGAGAACACCGTCATGTCTGTCAACATCTTCAAGCGCTTCTC 498
Db 241 AGTCTCTGTGTCTTGGGGCGAGAACACCGTCATGTCTGTCAACATCTTCAAGCGCTTCTC 300
QY 499 CCATGTCAACATCAAGCTGCGTCCCGGAGCGGAGGAGCGCCATCTTCAATGAGGTGGC 558
Db 301 CCATGTCAACATCAAGCTGCGTCCCGGAGCGGAGGAGCGCCATCTTCAATGAGGTGGC 360
QY 559 TCCAGGCTACTTCTCCCGGAGCGGCTCGCAGCTCCAGGTTTCAGGAGCGGTGGCAGAGCT 618
Db 361 TCCAGGCTACTTCTCCCGGAGCGGCTCGCAGCTCCAGGTTTCAGGAGCGGTGGCAGAGCT 420
QY 619 GGTGATCAAGAGCGCGCGGAGCTCCCATGTGGGGCTGTACATGTGGGACCTCTGTTGGGACA 678
Db 421 GGTGATCAAGAGCGCGCGGAGCTCCCATGTGGGGCTGTACATGTGGGACCTCTGTTGGGACA 480
QY 679 CCAGAGAAATCAACAGAAAGTCAAGTCAAGTGGAGGTTTCAGAGTGCAGAAACCCAGTCCGCC 738
Db 481 CCAGAGAAATCAACAGAAAGTCAAGTGGAGGTTTCAGTGCAGAAACCCAGTCCGCC 540
QY 739 TGACACTGGGTCTTGGGCTGTGCCAGCGGTGGTCACTGTCTTCTTCTTCTTCTTCTTCTG 798
Db 541 CGACACTGGGTCTTGGGCTGTGCCAGCGGTGGTCACTGTCTTCTTCTTCTTCTTCTTCTG 600
QY 799 TCTGTGCTATGTTGGCTGTGACAGTGGCTGTGTTCCAGCAAGCGCGGAGAGAGAGTT 858
Db 601 TCTGTGCTATGTTGGCTGTGACAGTGGCTGTGTTCCAGCAAGCGCGGAGAGAGAGTT 660
QY 859 CTTCCTCTTAGAAGCCCAAGATGAAGTCCAGAGCTCCAGAGCGGAGAGCGGAGCGGCT 918
Db 661 CTTCCTCTTAGAAGCCCAAGATGAAGTGGAGGCTTCAGAGCGGAGAGCGGAGCGGCT 720
QY 919 GAGCAGAGCGCTCCGTGAACTGTGAGCCCGCAGACTCCGAGGCCACCCCAAGCGCGCTGGC 978
Db 721 GAGCAGAGCGCTCCGTGAACTGTGAGCCCGCAGACTCCGAGGCCACCCCAAGCGCGCTGGC 780
QY 979 ACTGTGTTTCAAAACCTTCAACCTTGGAGCGCTTGGAGCTGTGTCTTCTTCTTCTTCTTGT 1038
Db 781 ACTGTGTTTCAAAACCTTCAACCTTGGAGCGCTTGGAGCTGTGTCTTCTTCTTCTTCTTGT 839
QY 1039 TTCCATATCCCGCAGACCCATAGCCCTTCGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1098
Db 840 TTCCATATCCCGCAGACCCATAGCCCTTCGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 899
QY 1099 CTGAGTGCAGACTTGGGTGGC- GGGGGCTGGGTCTCTCTGCTCCACCGGAGGAGAGAGAG 1157
Db 900 CTGAGTGCAGACTTGGGTGGGCGGGGCTGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 959

Db 615 GAGCAGAGCCTCGTGAACTGTGGACCCCGACACTCCGAGCCCAAGCCCGCTGGC 674

QY 979 ACTGGTGTCAACCTCACCCTTGGAGCCCTGGAGCTGTGTCTCCCGCCCAACCTTGT 1038

Db 675 ACTGGTGTCAACCTCACCCTTGGAGCCCTGGAGCTGTGTCTCCCGCCCAACCTTGT 733

QY 1039 TTCCATATGCCGACAGCCATAGCCGCTGCAAGCGCAGAGGACACAGGAGCCAGCC 1098

Db 734 TTCCATATGCCGACAGCCATAGCCGCTGCAAGCGCAGAGGACACAGGAGTAGCCACC 793

QY 1099 CTGAGTCCGACCTTGGGTGCG-GGGGCTGGGTCTCTGCTCCACCGGAGGACACAGA 1157

Db 794 CTGAGTCCGACCTTGGGTGCGGGGGCTGCTGCTCTCTGCTCCACCGGAGGACACAGA 853

QY 1158 CACCGGCTTGTGGCAGCGCTGGGCTCTGTGTCTAC 1193

Db 854 AGACACGGGCTTGTGGCAAGGCTTGGGGCTC 889

RESULT 13

ID ADB36353 standard; cDNA; 942 BP.

AC ADB36353;

XX

DT 04-DEC-2003 (first entry)

XX

DE Human immune response associated protein IRAP-8 cDNA SEQ ID NO:24.

XX

KW human; immune response associated protein; IRAP; anti-HIV; antiallergic;

KW antianemic; antidiabetic; antiarteriosclerotic; dermatological;

KW antinflammatory; antidiabetic; nephrotropic; antichyroid; thyromimetic;

KW immunosuppressive; antirheumatic; antidiabetic; osteopathic;

KW antibacterial; viricide; antiparasitic; protozoacide; fungicide;

KW cerebroprotective; neuroprotective; nootropic; antiparkinsonian;

KW antipsoriatic; cytostatic; cardiant; gene therapy;

KW immune system disorder; neurological disorder; developmental disorder;

KW muscle disorder; cell proliferative disorder; AIDS; allergy; anaemia;

KW asthma; atherosclerosis; contact dermatitis; Crohn's disease; diabetes;

KW glomerulonephritis; Grave's disease; Hashimoto's thyroiditis;

KW multiple sclerosis; rheumatoid arthritis; osteoporosis;

KW systemic lupus erythematosus; infection; stroke; Alzheimer's disease;

KW Parkinson's disease; psoriasis; cancer; cardiomyopathy; gene; ss.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT CDS 101..232

FT /*tag= a

FT /product= "IRAP-8"

FT

XX

FN WO2003074726-A2.

XX

PD 12-SEP-2003.

XX

PF 28-FEB-2003; 2003WO-US006307.

XX

PR 01-MAR-2002; 2002US-0361088P.

PR 27-MAR-2002; 2002US-0368494P.

PR 10-MAY-2002; 2002US-0379876P.

PR 28-JUN-2002; 2002US-0392641P.

XX

PA (INCY-) INCYTE GENOMICS INC.

XX

XX Chawla NK, Becha SD, Jin P, Tran UK, Kable AE, Lee S;

PI Marquis JP, Emerling BW, Ison CH, Ramkumar J, Chang H;

XX

DR WPI; 2003-722079/68.

DR P-PSDB; ADB36337.

XX

XX New human immune response associated proteins and polynucleotides, useful

PT for diagnosing, preventing or treating diseases or conditions associated

PT with aberrant enzyme expression, e.g. cancer, AIDS, atherosclerosis or

PT stroke.

XX Claim 5; Page 153; 158pp; English.

PS

XX ADB36346 to ADB36361 encode the human immune response associated proteins

CC given in ADB36330 to ADB36345, designated IRAP-1 to IRAP-16. IRAP

CC sequences have anti-HIV, antiallergic, antianemic, antidiabetic,

CC antiarteriosclerotic, dermatological, antinflammatory, antidiabetic,

CC nephrotropic, antithyroid, thyromimetic, immunosuppressive,

CC antirheumatic, antidiabetic, osteopathic, antibacterial, viricide,

CC antipsoriatic, protozoacide, fungicide, cerebroprotective,

CC neuroprotective, nootropic, antiparkinsonian, antipsoriatic, cytostatic

CC and cardiant activities, and can be used in gene therapy. The IRAP

CC proteins and polynucleotides can be used in diagnosing, preventing or

CC treating diseases or conditions associated with the decreased expression

CC or overexpression of IRAP, such as immune system, neurological,

CC developmental, muscle or cell proliferative disorders. The disorders may

CC include AIDS, allergies, anaemia, asthma, atherosclerosis, contact

CC dermatitis, Crohn's disease, diabetes, glomerulonephritis, Grave's

CC disease, Hashimoto's thyroiditis, multiple sclerosis, rheumatoid

CC arthritis, osteoporosis, systemic lupus erythematosus, infections (e.g.

CC bacterial, viral, fungal, parasitic or protozoal), stroke, Alzheimer's

CC disease, Parkinson's disease, psoriasis, cancer and cardiomyopathy. They

CC are also useful in assessing the effects of exogenous compounds on the

CC expression of nucleic acid and amino acid sequences of IRAP. The IRAP or

CC agonist or antagonist of the polypeptides, or in altering the expression

CC of the target polynucleotide and compounds that specifically bind to or

CC modulate the activity of the polypeptide.

XX

XX Sequence 942 BP; 175 A; 310 C; 282 G; 175 T; 0 U; 0 Other;

QY

Query Match 30.9%; Score 673.8; DB 10; Length 942;

Best Local Similarity 85.1%; Pred. No. 2.3e-144;

Matches 835; Conservative 0; Mismatches 37; Indels 109; Gaps 3;

QY 214 GCGCGGAGAGCTGCATCCCGAGGAGCGCTCCAGAGCGGAGCCCGGAGTGTTCAA 273

Db 1 GCGCGGAGAGCTGCATCCCGAGGAGCGCTCCAGAGCGGAGCCCGGAGTGTTCAA 60

QY 274 GAGCAGTGAACAGGACAGGAGCGGCGCCCAAGTCCACCGACCATGAGCCCTGCCCTGGC 333

Db 61 GAGCAGTGAACAGGACAGGAGCGGCGCCCAAGTCCACCGACCATGAGCCCTGCCCTGGC 120

QY 334 ATTCCCTGGCAGCGTTTCCAGGCGCTTGGGACCTCTGTTTGGTGGCTCTCTTGG 393

Db 121 ATTCCCTGGCAGCGTTTCCAGGCGCTTGGGACCTCTGTTTGGTGGCTCTCTTGG 180

QY 394 TGCTCAGAAATGAAGGCTGGGACAGCCCATCTGCACAGAGGGGGTAGTCTCTGTGTCFTG 453

Db 181 TGCTCAGAAATG-----

QY 454 GGGCGAGAAACACCGTCTATGTCCTGCAACATCTCCAAAGCCTTCTCCCATGTCAACATCAA 513

Db 192 -----AA 193

QY 514 GCTGGTGGCCACGGGCGAGGAGCGCCATCTTCAATGAGGTGGCTCCAGGCTACTTCTC 573

Db 194 GCTGGTGGCCACGGGCGAGGAGCGCCATCTTCAATGAGGTGGCTCCAGGCTACTTCTC 253

QY 574 CCGGAGACGGCTGGCAGCTCCAGGTTCCAGGAGCGCTGGGACAGCTGGTGGTATCAAGGGCGC 633

Db 254 CCGGAGACGGCTGGCAGCTCCAGGTTCCAGGAGCGCTGGGACAGCTGGTGGTATCAAGGGCGC 313

QY 634 CCGGAGCTCCCATGTGGGCTGTACATGTGGGACCTCTGTTGGGACACCGAGAGAAATAACAG 693

Db 314 CCGGAGCTCCCATGTGGGCTGTACATGTGGGACCTCTGTTGGGACACCGAGAGAAATAACAG 373

QY 694 ACAAGTCACGCTGAGGTTTTCAGGTCCAGAACCCAGTCCCGCCCTGTACACCTGGTTCG 753

Db 374 ACAAGTCACGCTGAGGTTTTCAGGTCCAGAACCCAGTCCCGCCCTGTACACCTGGTTCG 433

QY 754 GCCTGTGCCAGCGGTGGTCACTGTCTCTTCTTCTTGGTGGCTCTGTGTCTGTTCG 813

Db 434 GCCTGTGCAGGGTGGTCACTGCTCTTCTATCTCTTGGTCTGCTGCTATGTCGC 493
Qy 814 CTGTGTACAGGTCCGCTGTTCCAGCAACCGCGGAGAGAAAGTCTTCTCTCTAGAAC 873
Db 494 CTGGTACAGGTCCGCTGTTCCAGCAACCGCGGAGAGAAAGTCTTCTCTCTAGAAC 553
Qy 874 CCAGATGAAGTTCGAGCCCTCAGAGCGGAGAGCCAGAGGCGCTGAGCAGAGCTCCGC 933
Db 554 CCAGATGAAGTTCGAGCCCTCAGAGCGGAGAGCCAGAGGCGCTGAGCAGAGCTCCGC 613
Qy 934 TGAATCTGGACCCAGACTCCGAGCCACCCAGGCGCTGAGCAGAGCTGTTCAAAC 993
Db 614 TGAATCTGGACCCAGACTCCGAGCCACCCAGGCGCTGAGCAGAGCTGTTCAAAC 673
Qy 994 CTCACCACTTGGAGCCCTGGAGCTGCTGTCCTCCCTCCCAACCTTGTTCATATGCGCAG 1053
Db 674 CTCACCACTTGGAGCCCTGGAGCTGCTGTCCTCCCTCCCAACCTTGTTCATATGCGCAG 732
Qy 1054 ACCATAGCCGCTCAAGAGCAGAGACACAGAGAGCCAGCCCTGAGTGCAGACCTT 1113
Db 733 ACCATAGCCGCTCAAGAGAGAGAGACACAGAGTAGCCACCTGAGTGCAGACCTT 792
Qy 1114 GGTGCGC-GGGGCTGGTCTCTCTCCACCGGAGGCGACAGACCGGCTTCTGG 1172
Db 793 TGTGGCGGGGCGCTGGTCTCTCTCCACCGGAGGCGACAGACCGGCTTCTGG 852
Qy 1173 CAGGCTGGGCTCTCTGTCTAC 1193
Db 853 CTTCGCAAGCTTGGGGCTC 873

RESULT 14

AAL60549

ID AAL60549 standard; cDNA; 930 BP.

AC AAL60549;

XX 03-SEP-2003 (first entry)

XX Human organelle-associated protein (ORGA)-9 cDNA.

XX Human; organelle-associated protein; ORGA; cell proliferative disorder;

KW actinic keratosis; arteriosclerosis; reproductive disorder; infertility;

KW endometriosis; gastrointestinal disorder; ectopic pregnancy; hepatitis;

KW gastroenteritis; neurological disorder; Alzheimer's disease; cirrhosis;

KW glomerulonephritis; renal amyloidosis; renal failure; Addison's disease;

KW renal disorder; ovulatory defect; teratogenesis; Pick's disease; cancer;

KW Huntington's disease; urologic disorder; cystitis; dementia; dysphagia;

KW indigestion; gastritis; hypertension; gene therapy; bursitis; anorexia;

KW infection; urethritis; prostatitis; gene; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

FT 116..604

FT /*tag= a

FT /product= "Human ORGA protein"

XX WO200304171-A2.

XX 30-MAY-2003.

XX 15-NOV-2002; 2002WO-US036807.

XX 16-NOV-2001; 2001US-0332384P.

XX 13-DEC-2001; 2001US-0341187P.

XX 23-JAN-2002; 2002US-0351151P.

XX 27-FEB-2002; 2002US-0360269P.

XX 05-APR-2002; 2002US-0370637P.

XX 14-JUN-2002; 2002US-0388946P.

XX (INCY-) INCYTE GENOMICS INC.

XX Ramkumar J, Burford N, Xu Y, Chawla NK, Kable AE, Becha SD;
PI Richardson TW, Gorvad AE, Sprague WM, Elliott VS, Khare R;
PI Hawkins PR, Jin P, Lee SY, Marquis JP, Swarnakar A, Wilson AD;
XX Hafalia AJA, Bulloch S;
DR WPI; 2003-457603/43.
DR P-PSDB; AAO29896.
XX New organelle-associated proteins and polymucleotides, useful for
PT diagnosing, treating and/or preventing cell proliferative, reproductive,
PT gastrointestinal, neurological, urologic, and renal disorders.
XX Claim 5; Page 184; 194pp; English.

CC The invention relates to human organelle-associated proteins (ORGA) and
CC nucleic acid molecules encoding such proteins. ORGA sequences are useful
CC for diagnosing, treating and/or preventing cell proliferative disorders
CC (e.g. actinic keratosis, arteriosclerosis, bursitis, cirrhosis, cancers
CC or hepatitis), reproductive disorders (e.g. infertility, endometriosis,
CC ovulatory defect, ectopic pregnancy or teratogenesis), gastrointestinal
CC disorders (e.g. dysphagia, peptide oesophagitis, gastritis, indigestion,
CC anorexia or gastroenteritis), neurological disorders (e.g. Alzheimer's
CC disease, Pick's disease, Huntington's disease or dementia), urologic
CC disorders (e.g. infection, urethritis, cystitis, prostatitis or upper
CC tract infections) and renal disorders (e.g. glomerulonephritis, renal
CC amyloidosis, renal failure, Addison's disease or hypertension). The
CC invention is also useful in gene therapy. The present sequence is human
CC ORGA protein encoding cDNA

XX Sequence 930 BP; 182 A; 303 C; 275 G; 170 T; 0 U; 0 Other;

Query Match 29.2%; Score 636.4; DB 9; Length 930;
Best Local Similarity 82.7%; Pred. No. 8.6e-136;
Matches 824; Conservative 0; Mismatches 36; Indels 136; Gaps 3;

Qy 199 TTCTGGGGCTCCGGGGCGCGAGAGCTGCATCCAGAGGAGCGCTCCAGAGCGGAC 258

Db 1 TTCTGGGGCTCCGGGGCGCGAGAGCTGCATCCAGAGGAGCGCTCCAGAGCGGAC 60

Qy 259 CCGGAGGTGTTTCAAGAGCCAGTGCACAGGACCCAGGGGCCCAAGTCCACAGCCATGCA 318

Db 61 CCGGAGGTGTTTCAAGAGCCAGTGCACAGGACCCAGGGGCCCAAGTCCACAGCCATGCA 120

Qy 319 GACTGCCCCCTGGCATTCCTCTGGCCAGCTTTCAGGCGCCCTTGGGACCTCTCTGTTTT 378

Db 121 GACTGCCCCCTGGCATTCCTCTGGCCAGCTTTCAGGCGCCCTTGGGACCTCTCTGTTTT 180

Qy 379 GGCTGCTCTCTGAGTGTCTCAGAAATGAAGGTGGGACAGCCCAATCTGCACAGAGGGGT 438

Db 181 GGCTGCTCTCTGAGTGTCTCAGAAATGAAGGTGGGACAGCCCAATCTGCACAGAGGGGT 240

Qy 439 AGTCTCTGTGTCTGGGGCGAGAACACCGTGCATGTCTTGCACATCTTCAACGCCCTTCTC 498

Db 241 AGTCTCTGTGTCTGGGGCGAGAACACCGTGCATGTCTTGCACATCTTCAACGCCCTTCTC 300

Qy 499 CCATGTCAACATCAAGCTGCGTCCCGGAGAGAGCGCCATCTTCAATAGAGTGGC 558

Db 301 CCATGTCAACATCAAGCTGCGTCCCGGAGAGAGAGCGCCATCTTCAATAGAGTGGC 360

Qy 559 TCCAGGCTACTTCTCCCGGAGCGGCTGGCAGCTCCAGGTTTCAGGAGAGCGTGGCAGAGT 618

Db 361 TCCAGGCTACTTCTCCCGGAGCGGCTGGCAGCTCCAGGTTTCAGGAGAGCGTGGCAGAGT 420

Qy 619 GGTGATCAAGCGCGCGGGGACTCCCATGCTGGGCTGTACATGTGGGACCTCTGTTGGACA 678

Db 421 GGTGATCAAGCGCGCGGGGACTCCCATGCTGGGCTGTACATGTGGGACCTCTGTTGGACA 480

Qy 679 CCAGAGAAATAACAGACAAAGTCAAGTGGAGGTTTCAGGTTGCAGAAACCCAGTCCGCCCC 738

Db 481 CCAGAGAAATAACAGACAAAGTCAAGTGGAGGTTTCAGGTTGCAGAAACCCAGTCCGCCCC 516

Qy 739 TGACACTGGGTTCTGGGCTGTGCCAGCGGTGGTCACTGCTGTCTTCTCTCTTGGTGGC 798

Db 517 ----- 516
799 TCTGTCATGTTGGCTGTACAGTGGCTGTTCCAGCAACGCCGGGAGAGAGTT 858
Db 517 -----AGAGAGAGTT 526
859 CTTCTCTCTAGAACCCAGATGAAGTCCGAGCCCTCAGACGGGAGCCAGCAGGCGCT 918
Db 527 CTTCTCTCTAGAACCCAGATGAAGTCCGAGCCCTCAGACGGGAGCCAGCAGGCGCT 586
919 GAGCAGAGCCTCCGCTGAAGTGTGACCCAGATCCGAGCCCAACCAAGCCGCTGGC 978
Db 587 GAGCAGAGCCTCCGCTGAAGTGTGACCCAGATCCGAGCCCAACCAAGCCGCTGGC 646
979 ACTGCTGTCAACCCCTCAGATTCGAGCCCTGAGCTGTCTCTCCCTCCCAACCCCTTGT 1038
Db 647 ACTGCTGTCAACCCCTCAGATTCGAGCCCTGAGCTGTCTCTCCCTCCCAACCCCTTGT 705
1039 TTCCATATGCCGAGACCCATAGCCCTGCAAGGCAGAGGACACAGGAGCCAGCC 1098
Db 706 TTCCATATGCCGAGACCCATAGCCCTGCAAGGCAGAGGACACAGGAGCCAGCC 765
1099 CTGAGTCCGACCTTGGTGGC-GGGGCTGGGTCTCTCTCCCAACCCGAGGACACAGA 1157
Db 766 CTGAGTCCGACCTTGGTGGGCGGGGCTGGGTCTCTCTCCCAACCCGAGGACACAGA 825
1158 CACCGCTGTCTGGCAGGCTGGGCTCTGTGTAC 1193
Db 826 AGACACGGGCTTGTCTGGCAGGCTTGGGCGCTC 861

RESULT 15
ID AAL60550
XX AAL60550 standard; cDNA; 869 BP.
AC AAL60550;
XX
DT 03-SEP-2003 (first entry)
XX
DE Human organelle-associated protein (ORGA)-10 cDNA.
XX
KW Human; organelle-associated protein; ORGA; cell proliferative disorder;
KW actinic keratosis; arteriosclerosis; reproductive disorder; infertility;
KW endometriosis; gastrointestinal disorder; ectopic pregnancy; hepatitis;
KW gastroenteritis; neurological disorder; Alzheimer's disease; cirrhosis;
KW glomerulonephritis; renal amyloidosis; renal failure; Addison's disease;
KW renal disorder; ovulatory defect; teratogenesis; Pick's disease; cancer;
KW Huntington's disease; urologic disorder; cystitis; dementia; dysphagia;
KW indigestion; gastritis; hypertension; gene therapy; bursitis; anorexia;
KW infection; urethritis; prostatitis; gene; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 116..667
FT /*tag= a
FT /product= "Human ORGA protein"
XX
XX WO2003044171-A2.
XX
XX 30-MAY-2003.
XX
XX 15-NOV-2002; 2002WO-US036807.
XX
XX 16-NOV-2001; 2001US-0332384P.
XX 13-DEC-2001; 2001US-0341187P.
XX 23-JAN-2002; 2002US-0351151P.
XX 27-FEB-2002; 2002US-0360269P.
XX 05-APR-2002; 2002US-0370637P.
XX 14-JUN-2002; 2002US-0388946P.
XX
XX (INCY-) INCYTE GENOMICS INC.

XX
PI Ramkumar J, Burford N, Xu Y, Chawla NK, Kable AE, Becha SD;
PI Richardson TW, Gervad AE, Sprague WW, Elliott VS, Khare R;
PI Hawkins PR, Jin P, Lee SY, Marquis JP, Swarnakar A, Wilson AD;
PI Hafalia AJA, Bulloch S;
XX
XX WPI: 2003-457603/43.
XX P-PSDB; AAO29897.
XX
XX New organelle-associated proteins and polynucleotides, useful for
XX diagnosing, treating and/or preventing cell proliferative, reproductive,
XX gastrointestinal, neurological, urologic, and renal disorders.
XX
XX Claim 5; Page 184-185; 194pp; English.
XX
XX The invention relates to human organelle-associated proteins (ORGA) and
XX nucleic acid molecules encoding such proteins. ORGA sequences are useful
XX for diagnosing, treating and/or preventing cell proliferative disorders
XX (e.g. actinic keratosis, arteriosclerosis, bursitis, cirrhosis, cancers
XX or hepatitis), reproductive disorders (e.g. infertility, endometriosis,
XX or ovulatory defects, ectopic pregnancy or teratogenesis), gastrointestinal,
XX disorders (e.g. dysphagia, peptidic oesophagitis, gastritis, indigestion,
XX anorexia or gastroenteritis), neurological disorders (e.g. Alzheimer's
XX disease, Pick's disease, Huntington's disease or dementia), urologic
XX disorders (e.g. infection, urethritis, cystitis, prostatitis or upper
XX tract infections) and renal disorders (e.g. glomerulonephritis, renal
XX amyloidosis, renal failure, Addison's disease or hypertension). The
XX invention is also useful in gene therapy. The present sequence is human
XX ORGA protein encoding cDNA
XX
XX Sequence 869 BP; 161 A; 289 C; 262 G; 157 T; 0 U; 0 Other;
XX
XX Query Match 23.5%; Score 512.8; DB 9; Length 869;
XX Best Local Similarity 76.5%; Pred. No. 1.9e-107;
XX Matches 762; Conservative 0; Mismatches 37; Indels 197; Gaps 3;
XX
QY 199 TTCTGGGGCTCCGGGGCGGGAAGCTGCATCCAGAGAGCGCTCCAGGAGCGGAC 258
Db 1 TTCTGGGGCTCCGGGGCGGGAAGCTGCATCCAGAGAGCGCTCCAGGAGCGGAC 60
QY 259 CCGGAGTGTTCAGAGAGCCAGTCAAGAACAGGCGGCCAGTCCACAGCCATGCA 318
Db 61 CCGGAGTGTTCAGAGAGCCAGTCAAGAACAGGCGGCCAGTCCACAGCCATGCA 120
QY 319 GACCTGCCCTTGGCATTCCTGCGCCAGTTCCTCCAGGCGCTTCCTCTGTTTTT 378
Db 121 GACCTGCCCTTGGCATTCCTGCGCCAGTTCCTCCAGGCGCTTCCTCTGTTTTT 180
QY 379 GGCTGCCCTTGTAGTGTCTCAGAAATGAAGGCTGGGACAGCCCATCTGCAAGAGGGGT 438
Db 181 GGCTGCCCTTGTAGTGTCTCAGAAATGAAGGCTGGGACAGCCCATCTGCAAGAGGGGT 240
QY 439 AGTCTCTGTCTTGGGGGAGAACACCGTCAATGCTTCCCAACATCTCCAAGCCCTCTC 498
Db 241 AGTCTCTGTCTTGGGGGAGAACACCGTCAATGCTTCCCAACATCTCCAAGCCCTCTC 300
QY 499 CCATGTCAACATCAAGCTGCGTCCCAAGGAGAGAGCGCCATCTTCAATAGGTTGC 558
Db 301 CCATGTCAACATCAAGCTGCGTCCCAAGGAGAGAGCGCCATCTTCAATAGGTTGC 360
QY 559 TCCAGGCTACTTCTCCCGGGAAGGCTGCGAGCTCCAGTTTTCAGGAGGCGTGGCAGCT 618
Db 361 TCCAGGCTACTTCTCCCGGGAAGGCTGCGAGCTCCAGTTTTCAGGAGGCGTGGCAGCT 420
QY 619 GGTGATCAAGGGGCGCGGAGCTCCCATGCTGGGTGTACATGTGGCAGCTCTGTGGGACA 678
Db 421 GGTGATCAAGGGGCGCGGAGCTCCCATGCTGGGTGTACATGTGGCAGCTCTGTGGGACA 480
QY 679 CCAGAGAAATACAGACAAGTCAAGTGGAGGTTTCAGGTGAGAACCCAGCTCCGCCCC 738
Db 481 ----- 480
QY 739 TGACACTGGGTTCTGGCTGTGCCAGCGGTGGTCACTGCTGTCTTCTTCTTGTGTCG 798

```
Db 481 ----- 480
Qy 799 TCTGTCATGTTTCGCTGTGTACAGTGCCTGTTCCAGCAACGCGGGAGAGAAGTT 858
Db 481 ----- 480
Qy 859 CTTCTCTAGAACCCAGATGAAGTGCAGCCCTCAGAGCGGAGCCAGCAGGGCT 918
Db 481 ----- 480
Qy 919 GAGCAGAGCTCCGCTGAACTGTGACCCAGACTCCGAGCCCAAGGCGCTGGC 978
Db 526 GAGCAGAGCTCCGCTGAACTGTGACCCAGACTCCGAGCCCAAGGCGCTGGC 585
Qy 979 ACTGCTGTTCAAACCTCACCATTGGAGCCCTGGAGCTGCTGTCCCGCCCAACCTTGT 1038
Db 586 ACTGCTGTTCAAACCTCACCATTGGAGCCCTGGAGCTGCTGT-CCCCCAACCTTGT 644
Qy 1039 TTCCATATGCCCGCAGACCCATAGCCGCTGCAAGGAGAGAGGACACAGGAGAGCCAGCC 1098
Db 645 TTCCATATGCCCGCAGACCCATAGCCGCTGCAAGGAGAGAGGACACAGGAGTAGCCACC 704
Qy 1099 CTGAGTCCGACCTTGGGTGGC-GGGGCTGGGTCTCTGTCACCCGAGGGCACAGA 1157
Db 705 CTGAGTCCGACCTTGGGTGGGCGGGGCTGGGTCTCTGTCACCCGAGGGCACACA 764
Qy 1158 CACCGGCTTGTGGCAGGCTGGGCTCTGTGTAC 1193
Db 765 AGACACCGGGCTTGTGGCAAGGCTTGGGCGCTC 800
```

Search completed: March 20, 2005, 02:10:42

Job time : 1195 secs

THIS PAGE BLANK (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 20, 2005, 01:40:02 ; Search time 380 Seconds
(without alignments)

9387.061 Million cell updates/sec

Title: US-10-080-522-2

Perfect score: 2180

Sequence: 1 ATTCCTGCTTCCTTAGCGT.....TTAAATAAGTTCGTCG 2180

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*

1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2180	100.0	2180	2	US-08-755-559-2
2	2180	100.0	2180	3	US-09-210-474-2
3	2180	100.0	2180	3	US-09-539-774-2
4	1951	89.5	2000	4	US-09-997-165-3
5	1951	89.5	2000	4	US-09-949-016-426
6	1889.6	86.7	2001	4	US-09-949-016-2916
7	1241.6	57.0	16738	4	US-09-949-016-12168
8	1241.6	57.0	16738	4	US-09-949-016-14678
9	566	26.0	601	4	US-09-949-016-26332
10	566	26.0	601	4	US-09-949-016-26333
11	541.4	24.8	601	4	US-09-949-016-105202
12	541.4	24.8	601	4	US-09-949-016-26334
13	155.4	7.1	1112	4	US-09-949-016-105203
14	138.6	6.4	601	4	US-09-997-165-7
15	138.6	6.4	601	4	US-09-949-016-26332
16	55	2.5	7218	1	US-09-949-016-105201
17	48.4	2.2	19203	4	US-08-232-463-14
18	48.4	2.2	19210	4	US-08-232-463-14
19	48	2.2	19210	4	US-09-949-016-15519
20	48	2.2	31868	4	US-09-949-016-12129
21	46.2	2.1	4112	1	US-09-949-016-11907
22	46.2	2.1	4112	1	US-08-340-203A-2
23	46.2	2.1	4112	2	US-08-452-567-2
24	46.2	2.1	4112	3	US-08-452-567-2
25	46.2	2.1	4616	1	US-09-085-407-2
26	46.2	2.1	4616	2	US-08-340-203A-1
27	46.2	2.1	4616	2	US-08-452-567-1

c 28	46.2	2.1	4616	3	US-09-085-407-1	Sequence 1, Appli
c 29	45.8	2.1	61158	4	US-09-949-016-15041	Sequence 15041, A
c 30	45.8	2.1	75431	4	US-09-949-016-15122	Sequence 15122, A
c 31	45.8	2.1	670689	4	US-09-949-016-12505	Sequence 12505, A
c 32	45.8	2.1	670690	4	US-09-949-016-14207	Sequence 14207, A
c 33	45.4	2.1	3561	4	US-09-616-289-48	Sequence 48, Appl
c 34	45.2	2.1	3120	4	US-09-016-434-1050	Sequence 1050, Ap
c 35	45	2.1	46244	4	US-09-949-016-13508	Sequence 13508, A
c 36	44.8	2.1	6029	4	US-09-949-016-13342	Sequence 13342, A
c 37	44.8	2.1	11093	4	US-09-949-016-12142	Sequence 12142, A
c 38	44.6	2.0	1746	4	US-09-573-080A-186	Sequence 186, App
c 39	44.6	2.0	152331	3	US-09-128-155-16	Sequence 16, Appl
c 40	44.2	2.0	154746	4	US-09-827-688-8	Sequence 8, Appli
c 41	44.2	2.0	154746	4	US-09-827-688-8	Sequence 8, Appli
c 42	44	2.0	2823	1	US-08-398-008A-1	Sequence 1, Appli
c 43	44	2.0	2823	2	US-08-893-333-1	Sequence 1, Appli
c 44	44	2.0	31391	4	US-09-949-016-14319	Sequence 14319, A
c 45	43.8	2.0	430	4	US-09-621-976-16656	Sequence 16656, A

ALIGNMENTS

RESULT 1

US-08-755-559-2
Sequence 2, Application US/08755559
Patent No. 5912142
GENERAL INFORMATION:
APPLICANT: KAUFMAN, RUSSEL E.
APPLICANT: SLENTZ-KESLER, KIMBERLY
TITLE OF INVENTION: GENE PRODUCT OVER EXPRESSED IN CANCER
TITLE OF INVENTION: CELLS
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHIVE P.C.
STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/755,559
FILING DATE: 22-NOV-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: WILSON, MARY J.
REGISTRATION NUMBER: 32,955
REFERENCE/DOCKET NUMBER: 1579-116
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2180 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-755-559-2

Query Match 100.0%; Score 2180; DB 2; Length 2180;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATTCCTGCTTCCTTAGCGTGAACCGGGTGGGTCCTCCCGTGAATAAATTAATTCAC 60
DB 1 ATTCCTGCTTCCTTAGCGTGAACCGGGTGGGTCCTCCCGTGAATAAATTAATTCAC 60

QY	61	CGTACGCTTGTGTGAACGGGGTGGTTCCGAAACTTGGAGGCTTCCGTAACCCAG	120	1141	CACCCGAGGGGACACAGACCCGCTTGTGGAGGCTGGGCTCTGTGTACCCACTCC	1200
Db	61	CGTACGCTTGTGTGAACGGGGTGGTTCCGAAACTTGGAGGCTTCCGTAACCCAG	120	1201	TGGGTGGGTGACAGACCCCTTCCCTCCACCCCGGAGTCTTCCAGGCTTCTCTCACT	1260
QY	121	CTCCTTCTCTCACTGGGAGTGGTCCCGGGGGTCCCGGCTCTCTCTCTCTCTCTCT	180	1201	TGGGTGGGTGACAGACCCCTTCCCTCCACCCCGGAGTCTTCCAGGCTTCTCTCACT	1260
Db	121	CTCCTTCTCTCACTGGGAGTGGTCCCGGGGGTCCCGGCTCTCTCTCTCTCTCTCT	180	1261	TTCCAAATGGAAACCACTTCCCTCCACCCCGGAGTCTTCCAGGCTTCTCTCACT	1320
QY	181	CT	240	1261	TTCCAAATGGAAACCACTTCCCTCCACCCCGGAGTCTTCCAGGCTTCTCTCACT	1320
Db	181	CT	240	1321	CTCTGGCTCATCAAAACCCACAGACCCGAGTCTCTCTCTCTCTCTCTCTCTCTCT	1380
QY	241	GGCGGTCCAGGAGCGGACCCGGAGTGTTCAGAGCCAGTGACAAAGACAGGGGCCA	300	1321	CTCTGGCTCATCAAAACCCACAGACCCGAGTCTCTCTCTCTCTCTCTCTCTCTCT	1380
Db	241	GGCGGTCCAGGAGCGGACCCGGAGTGTTCAGAGCCAGTGACAAAGACAGGGGCCA	300	1381	CCCCAGGTGGGGTCCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1440
QY	301	AGTCCACAGAGCTGACAGACCTGCCCTGTCATCTCTCTCTCTCTCTCTCTCTCT	360	1381	CCCCAGGTGGGGTCCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1440
Db	301	AGTCCACAGAGCTGACAGACCTGCCCTGTCATCTCTCTCTCTCTCTCTCTCTCT	360	1441	CTGCGGAGCTCAGACACACTGGAGTTCAGGGCTGGGGGGGCTTGGCACAATCTGTC	1500
QY	361	TGGGACCT	420	1441	CTGCGGAGCTCAGACACACTGGAGTTCAGGGCTGGGGGGGCTTGGCACAATCTGTC	1500
Db	361	TGGGACCT	420	1501	CCTTGGCTATGAGCAGGCTTGGGGGCTTCCGCGCAGCCCGGGGGGCGAGGTAGGG	1560
QY	421	CATCTGCACAGAGGGGTAGTCTCTGTCTTGGGGGAGAACACCGTCACTCTCTGCAA	480	1501	CCTTGGCTATGAGCAGGCTTGGGGGCTTCCGCGCAGCCCGGGGGGCGAGGTAGGG	1560
Db	421	CATCTGCACAGAGGGGTAGTCTCTGTCTTGGGGGAGAACACCGTCACTCTCTGCAA	480	1561	TCTGGGGCTTAGAGGCTGGGATGGCTCTGGGCCCCACCCGCGAGGGGCGAAGCGCAGCC	1620
QY	481	CATCTGCACAGGGGTAGTCTCTGTCTTGGGGGAGAACACCGTCACTCTCTGCAA	540	1561	TCTGGGGCTTAGAGGCTGGGATGGCTCTGGGCCCCACCCGCGAGGGGCGAAGCGCAGCC	1620
Db	481	CATCTGCACAGGGGTAGTCTCTGTCTTGGGGGAGAACACCGTCACTCTCTGCAA	540	1621	GGGCTGGGAGCGCGCGGGCTCGGGCTGGGGGGTCAAGTGGACGCTGCGCTCCGGGG	1680
QY	541	CATCTGCACAGGGGTAGTCTCTGTCTTGGGGGAGAACACCGTCACTCTCTGCAA	600	1621	GGGCTGGGAGCGCGCGGGCTCGGGCTGGGGGGTCAAGTGGACGCTGCGCTCCGGGG	1680
Db	541	CATCTGCACAGGGGTAGTCTCTGTCTTGGGGGAGAACACCGTCACTCTCTGCAA	600	1681	CTGCTGGGAGCTTCTTGGAGCCAGATCTGTTCTATGCTTTTGTCTTGTCTGTCAT	1740
QY	601	GGGAGGCTGGGACAGCTGTGTATCAAGGGCGCGGGGCTCCCATGCTGGGCTGTACAT	660	1681	CTGCTGGGAGCTTCTTGGAGCCAGATCTGTTCTATGCTTTTGTCTTGTCTGTCAT	1740
Db	601	GGGAGGCTGGGACAGCTGTGTATCAAGGGCGCGGGGCTCCCATGCTGGGCTGTACAT	660	1741	CCTGCGGAGCTTCTTGGAGCCAGATCTGTTCTATGCTTTTGTCTTGTCTGTCAT	1800
QY	661	GTGGACCTCTGGGACACAGAGAAATAACAGCAAGTCACTGGAGGTTTCAGGTGC	720	1741	CCTGCGGAGCTTCTTGGAGCCAGATCTGTTCTATGCTTTTGTCTTGTCTGTCAT	1800
Db	661	GTGGACCTCTGGGACACAGAGAAATAACAGCAAGTCACTGGAGGTTTCAGGTGC	720	1801	GCCCTTTGATGTCCTTCTATGAGGCTGGGAGTCAAGTCAAGTCAAGTCAAGTCAAG	1860
QY	721	AGAAACCCAGTCCGCTGACACTGGGTTCTGGCTGTGCGAGCGGTGTCTCTCTCT	780	1801	GCCCTTTGATGTCCTTCTATGAGGCTGGGAGTCAAGTCAAGTCAAGTCAAGTCAAG	1860
Db	721	AGAAACCCAGTCCGCTGACACTGGGTTCTGGCTGTGCGAGCGGTGTCTCTCTCT	780	1861	TTTGAAGAGTTTCCATGACTCGAATATCTGAATGAAGAAACAAACCGACTCAAAACC	1920
QY	781	CTTCATCTCTTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	840	1861	TTTGAAGAGTTTCCATGACTCGAATATCTGAATGAAGAAACAAACCGACTCAAAACC	1920
Db	781	CTTCATCTCTTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	840	1921	TCCAAAGTAGCTCAAATGCAATTTTAAATGGAAGAAACAAATCTGAAGAAAGCTCTT	1980
QY	841	ACGCGGAGAGAGAGTTCTTCTCTAGAACCCAGATGAAGTCCGAGCCTCAGAGC	900	1921	TCCAAAGTAGCTCAAATGCAATTTTAAATGGAAGAAACAAATCTGAAGAAAGCTCTT	1980
Db	841	ACGCGGAGAGAGAGTTCTTCTCTAGAACCCAGATGAAGTCCGAGCCTCAGAGC	900	1981	TAGTGGCTTTAAGCCCAAAAGCTCCCTAAGGCTCTCGAGATGAAGACGGGGGGAGC	2040
QY	901	GGGAGCCAGAGGCGCTGAGCAGAGCTCCGCTGMACTGGAGCCCAAGACTCCGAGCC	960	1981	TAGTGGCTTTAAGCCCAAAAGCTCCCTAAGGCTCTCGAGATGAAGACGGGGGGAGC	2040
Db	901	GGGAGCCAGAGGCGCTGAGCAGAGCTCCGCTGMACTGGAGCCCAAGACTCCGAGCC	960	2041	CCGAGCCAGGTGGAGACCCCGCAGGACGCGGGGGGCTCGGTGACGAGGCTTCGACAG	2100
QY	961	CACCCCAAGGCGCTGGCACTGGTGTTCAAACCCCTCACACTTGGAGCCTTGGAGTGT	1020	2041	CCGAGCCAGGTGGAGACCCCGCAGGACGCGGGGGGCTCGGTGACGAGGCTTCGACAG	2100
Db	961	CACCCCAAGGCGCTGGCACTGGTGTTCAAACCCCTCACACTTGGAGCCTTGGAGTGT	1020	2101	CCGCGCGCTTGAAGGTGGGGTCCAGAGGGGCTTGAAGAGGGGCTTGTGTCTTCGGG	2160
QY	1021	GTCCCCCAACCCCTTGTTCATATGGCGAGACCCATAGCCGCTTCAGAGGAGAGAG	1080	2101	CCGCGCGCTTGAAGGTGGGGTCCAGAGGGGCTTGAAGAGGGGCTTGTGTCTTCGGG	2160
Db	1021	GTCCCCCAACCCCTTGTTCATATGGCGAGACCCATAGCCGCTTCAGAGGAGAGAG	1080	2161	TTAAATAAGGTTCCGTCGCG 2180	
QY	1081	GACAGGAGGAGCCCTGAGTCCGACCTTGGGTGGCGGGGCTTGGGTCTCTCTGTC	1140	2161	TTAAATAAGGTTCCGTCGCG 2180	
Db	1081	GACAGGAGGAGCCCTGAGTCCGACCTTGGGTGGCGGGGCTTGGGTCTCTCTGTC	1140			
QY	1141	CACCCGAGGGGACAGACACCGGCTTGTGGAGGCTGGGCTCTGTGTACCCACTCC	1200			

US-09-210-474-2
; Sequence 2, Application US/09210474
; Patent No. 6072034
; GENERAL INFORMATION:
; APPLICANT: KAUFMAN, RUSSEL E.
; APPLICANT: SLENTZ-KESLER, KIMBERLY
; TITLE OF INVENTION: GENE PRODUCT OVER EXPRESSED IN CANCER
; TITLE OF INVENTION: CELLS
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHUYE P.C.
; STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/210,474
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/755,559
; FILING DATE: 22-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: WILSON, MARY J.
; REGISTRATION NUMBER: 32,955
; REFERENCE/DOCKET NUMBER: 1579-116
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2180 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-09-210-474-2

Query Match 100.0%; Score 2180; DB 3; Length 2180;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATTCCTGCTCTTTAGCGTGAACGCGGGTGGCGTCCCGTGAATAATAAATTCAC 60
DB 1 ATTCCTGCTCTTTAGCGTGAACGCGGGTGGCGTCCCGTGAATAATAAATTCAC 60
QY 61 CGTCACGCTGTGTGAACGCGGGTGGTTCGCGAACTTGGAGGCTTCCCGTAACCCAG 120
DB 61 CGTCACGCTGTGTGAACGCGGGTGGTTCGCGAACTTGGAGGCTTCCCGTAACCCAG 120
QY 121 CTCCTTCTCTCATCTGGAGTGGGTCCCGCGGGTCCCGCCCTCTCTCTGGCCCTC 180
DB 121 CTCCTTCTCTCATCTGGAGTGGGTCCCGCGGGTCCCGCCCTCTCTCTGGCCCTC 180
QY 181 CCTCTCGTGTCTTTCAATTTCTTGGGGTCCGGGGCGGAGAGCTGCATCCAGAGGA 240
DB 181 CCTCTCGTGTCTTTCAATTTCTTGGGGTCCGGGGCGGAGAGCTGCATCCAGAGGA 240
QY 241 GCGGCTCCAGAGCGGACCGCGGAGTGTTCAGAGCCAGTCAAGAGCCAGGGGCCCA 300
DB 241 GCGGCTCCAGAGCGGACCGCGGAGTGTTCAGAGCCAGTCAAGAGCCAGGGGCCCA 300
QY 301 AGTCCACACAGCATGCACACTGCCCTCTGGCATTTCCCTGGCCACGTTTCCAGGCCCT 360
DB 301 AGTCCACACAGCATGCACACTGCCCTCTGGCATTTCCCTGGCCACGTTTCCAGGCCCT 360
QY 361 TGGGACCCCTCTCTGTTTTTGGGCTGCCTCTTTGAGTGTCTCAGAAATGAAGGCTGGGACAGCCC 420

DB 361 TGGGACCCCTCTCTGTTTTTGGGCTGCCTCTTTGAGTGTCTCAGAAATGAAGGCTGGGACAGCCC 420
QY 421 CATCTGCAACAGAGGGGTAGTCTCTGTGTCTTTGGGCGGAGAACACCGTCTATGCTCTGCAAA 480
DB 421 CATCTGCAACAGAGGGGTAGTCTCTGTGTCTTTGGGCGGAGAACACCGTCTATGCTCTGCAAA 480
QY 481 CATCTCCAAACGCTTTCTCCCATGTCAACATCAAGCTGCGTCCCAACGCGGACGAGAGCGC 540
DB 481 CATCTCCAAACGCTTTCTCCCATGTCAACATCAAGCTGCGTCCCAACGCGGACGAGAGCGC 540
QY 541 CATCTTCAATGAGGTGGCTCCAGGCTACTTCTCCGCGGACGCGTGGCAGCTCCAGGTTCA 600
DB 541 CATCTTCAATGAGGTGGCTCCAGGCTACTTCTCCGCGGACGCGTGGCAGCTCCAGGTTCA 600
QY 601 GGGAGCGGTGGCACAGCTGGTGATCAAAAGGCGCCCGGACTCCCATGCTGGGCTGTGACAT 660
DB 601 GGGAGCGGTGGCACAGCTGGTGATCAAAAGGCGCCCGGACTCCCATGCTGGGCTGTGACAT 660
QY 661 GTGGCACCTCGTGGGACACAGAGAAATAACAGACAAGTCAAGCTGGAGGTTTCAGGTGC 720
DB 661 GTGGCACCTCGTGGGACACAGAGAAATAACAGACAAGTCAAGCTGGAGGTTTCAGGTGC 720
QY 721 AGAACCCAGTCCGCGCTTGACACTGGGTTCTGGGCTGTGGCAGCGGTGCTCACTGCTGT 780
DB 721 AGAACCCAGTCCGCGCTTGACACTGGGTTCTGGGCTGTGGCAGCGGTGCTCACTGCTGT 780
QY 781 CTTTCATCTCTTGGTGGTCTGGTCTGATGTTCCCTTGGTACAGGTGCGGCTGTTCAGCA 840
DB 781 CTTTCATCTCTTGGTGGTCTGGTCTGATGTTCCCTTGGTACAGGTGCGGCTGTTCAGCA 840
QY 841 AGCGCGGAGAGAAAGTTCTTCTCTAGAACCCAGATGAAGGTGCGACGCTCCAGAGC 900
DB 841 AGCGCGGAGAGAAAGTTCTTCTCTAGAACCCAGATGAAGGTGCGACGCTCCAGAGC 900
QY 901 GGGAGCCAGACAGGCGCTGAGCAGAGCTCCGCTGAACTGTGTGACCCAGACTCCGAGCC 960
DB 901 GGGAGCCAGACAGGCGCTGAGCAGAGCTCCGCTGAACTGTGTGACCCAGACTCCGAGCC 960
QY 961 CACCCAAAGGCGCTGGCACTGGTGTCAAAACCCCTCACACCTTGGAGGCTTGGAGTGT 1020
DB 961 CACCCAAAGGCGCTGGCACTGGTGTCAAAACCCCTCACACCTTGGAGGCTTGGAGTGT 1020
QY 1021 GTCCCCCCTTGTTCATATGCGCAGACCCATAGCCGCTTGAAGGCGAGAG 1080
DB 1021 GTCCCCCCTTGTTCATATGCGCAGACCCATAGCCGCTTGAAGGCGAGAG 1080
QY 1081 GACACAGAGAGCAGCCCTGAGTCCGACCTTGGGTGGCGGGGCTTGGGTCTCTGCTCC 1140
DB 1081 GACACAGAGAGCAGCCCTGAGTCCGACCTTGGGTGGCGGGGCTTGGGTCTCTGCTCC 1140
QY 1141 CACCGGAGGCGCACAGACACGCGCTTGTGGCAGGCTGGGCTCTGTGTCAACCACTCC 1200
DB 1141 CACCGGAGGCGCACAGACACGCGCTTGTGGCAGGCTGGGCTCTGTGTCAACCACTCC 1200
QY 1201 TGGGTGCGTGGCAGACCTTCCCTCCACCCCGCAGGTCTTCAAGCTCTGTCTCTCACT 1260
DB 1201 TGGGTGCGTGGCAGACCTTCCCTCCACCCCGCAGGTCTTCAAGCTCTGTCTCTCACT 1260
QY 1261 TTCCAAAATGGAACCACTCCCGCAGACCCGACTTACAGAGACGATGCCCTCC 1320
DB 1261 TTCCAAAATGGAACCACTCCCGCAGACCCGACTTACAGAGACGATGCCCTCC 1320
QY 1321 CTCTGCCCTCATCAAAACCCAGACCTCCCTTCTGCGCCACCCAGGCTGGTCCGG 1380
DB 1321 CTCTGCCCTCATCAAAACCCAGACCTCCCTTCTGCGCCACCCAGGCTGGTCCGG 1380
QY 1381 CCCCAGGTGTGGGGTCCGCTCTCTCACTCCAGGCGCTCCGCGCCCAAGTGAAGGGGCC 1440
DB 1381 CCCCAGGTGTGGGGTCCGCTCTCTCACTCCAGGCGCTCCGCGCCCAAGTGAAGGGGCC 1440
QY 1441 CTGCGGAGCCTCAGACACACTGAGGATTCAGGGCTGGGGGCGCTTGGCAGATACCTGTTC 1500

1441 CTGCGGAGCCTCAGACACACTGGAGTTCAAGGCTGGGGGGCCCTTGGCAATACCTGTC 1500
1501 CTTTGGCTATAGCAGGCTTTGGGGCCCTTCCGGCAGCCCGGGGGCCGAGTAGGG 1560
1501 CTTTGGCTATAGCAGGCTTTGGGGCCCTTCCGGCAGCCCGGGGGCCGAGTAGGG 1560
1561 TCTGGGGCTTAGAGGCTGGAGTGTCTCTGGCCCAACCGCAGGGGGCAAGCCAGGCC 1620
1561 TCTGGGGCTTAGAGGCTGGAGTGTCTCTGGCCCAACCGCAGGGGGCAAGCCAGGCC 1620
1621 GGGCTGGAGGGCGGGGGCGGCTCGGGCTGGGGGGTCAAGTGACCTGCCTCGGGGG 1680
1621 GGGCTGGAGGGCGGGGGCGGCTCGGGCTGGGGGGTCAAGTGACCTGCCTCGGGGG 1680
1681 CTGGTGGCGCATCCCTCAGTCCCTCGGCCACCCGGGGGTGCTCCTCTGTGCCCCACCA 1740
1681 CTGGTGGCGCATCCCTCAGTCCCTCGGCCACCCGGGGGTGCTCCTCTGTGCCCCACCA 1740
1741 CTGGGAGCCTTTTGGACCCAGATCTGTTCATGCTTTTGTCTTGTCTTGTCTTGTCTTGTCT 1800
1741 CTGGGAGCCTTTTGGACCCAGATCTGTTCATGCTTTTGTCTTGTCTTGTCTTGTCTTGTCT 1800
1801 GCCCTTGTATGCTTTCATCTGTATGGGTGGAAATCACCAGGGAATCCCCCTTCAGTTC 1860
1801 GCCCTTGTATGCTTTCATCTGTATGGGTGGAAATCACCAGGGAATCCCCCTTCAGTTC 1860
1861 TTTGAAAAAGTTCCATGACTCGAATATCTGAATGAAGAAAAAACAACCGACTCACAAACC 1920
1861 TTTGAAAAAGTTCCATGACTCGAATATCTGAATGAAGAAAAAACAACCGACTCACAAACC 1920
1921 TCCAAGTAGCTCCAAATGCANTTTTTAAATGGAACAAAAAATCTGAAAGAAACGTCTT 1980
1921 TCCAAGTAGCTCCAAATGCANTTTTTAAATGGAACAAAAAATCTGAAAGAAACGTCTT 1980
1981 TAGTGGCTTTAAGCCCCCAAAACGCTCCTAAGGGCTCTCGAGATGAAGACGGGGGGAGC 2040
1981 TAGTGGCTTTAAGCCCCCAAAACGCTCCTAAGGGCTCTCGAGATGAAGACGGGGGGAGC 2040
2041 CCCAGCAGGTGAGACCCCGCAGGACCGGGCGGCGCCCGGTGACCGAGGCTTCGCACAG 2100
2041 CCCAGCAGGTGAGACCCCGCAGGACCGGGCGGCGCCCGGTGACCGAGGCTTCGCACAG 2100
2101 CCGGCGCCCTGAGGGTGGGGCGGAGCCAGGCTTCAAGAGGGCGGGTGTGTCTTCGGG 2160
2101 CCGGCGCCCTGAGGGTGGGGCGGAGCCAGGCTTCAAGAGGGCGGGTGTGTGTCTTCGGG 2160
2161 TTAATAAGGTTCCGTCG 2180
2161 TTAATAAGGTTCCGTCG 2180

RESULT 3
US-09-539-774-2
; Sequence 2, Application US/09539774
; Patent No. 6350615
; GENERAL INFORMATION:
; APPLICANT: KAUFMAN, RUSSEL E.
; APPLICANT: SLENTZ-KESLER, KIMBERLY
; TITLE OF INVENTION: GENE PRODUCT OVER EXPRESSED IN CANCER
; TITLE OF INVENTION: CELLS
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P. C.
; STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/539,774
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/210,474
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: WILSON, MARY J.
REGISTRATION NUMBER: 32,955
REFERENCE/DOCKET NUMBER: 1579-116
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ. ID NO. 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2180 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-09-539-774-2
Query Match 100.0%; Score 2180; DB 3; Length 2180;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATTCCTGCTTCTTTAGCGTGAACCGGGTGGCGTGCCTCCCGTGAATAATAAATTCAC 60
DB 1 ATTCCTGCTTCTTTAGCGTGAACCGGGTGGCGTGCCTCCCGTGAATAATAAATTCAC 60
QY 61 CGTCAACGCTTCTTGTGAACCGGGTGGTTCGCAAACTTGGAGGCTTCCCGTAAACCCAG 120
DB 61 CGTCAACGCTTCTTGTGAACCGGGTGGTTCGCAAACTTGGAGGCTTCCCGTAAACCCAG 120
QY 121 CTCTTCTCTATCTGGAGGTTGGTCCCGCGGGGTCCCGCGCTCTCTCTGCGCCCTC 180
DB 121 CTCTTCTCTATCTGGAGGTTGGTCCCGCGGGGTCCCGCGCTCTCTCTGCGCCCTC 180
QY 181 CTTCTCTGCTCTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 240
DB 181 CTTCTCTGCTCTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 240
QY 241 GCGGCTCAGAGAGCGGACCCCGGAGTGTTCAGAGCCAGTCAAGAGCCAGGAGGAGCCCA 300
DB 241 GCGGCTCAGAGAGCGGACCCCGGAGTGTTCAGAGCCAGTCAAGAGCCAGGAGGAGCCCA 300
QY 301 AGTCCCAACGACATGACAGCCTGCGCCCTCTGGCATTTCCCTGGCCACGTTTCCAGGCCCT 360
DB 301 AGTCCCAACGACATGACAGCCTGCGCCCTCTGGCATTTCCCTGGCCACGTTTCCAGGCCCT 360
QY 361 TGGGACCCCTCTTGTGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 420
DB 361 TGGGACCCCTCTTGTGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 420
QY 421 CATCTGCACAGAGGGGTAGTCTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 480
DB 421 CATCTGCACAGAGGGGTAGTCTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 480
QY 481 CATCTCCAAACGCTTCTCCCATGTCAACATCAAGCTGCGTCCCAAGAGGAGAGCCG 540
DB 481 CATCTCCAAACGCTTCTCCCATGTCAACATCAAGCTGCGTCCCAAGAGGAGAGCCG 540
QY 541 CATCTTCAATGAGGTGCTTCCAGGCTACTTCTTCCCGGGGACCGCTGCGAGCTCCAGGTTCA 600
DB 541 CATCTTCAATGAGGTGCTTCCAGGCTACTTCTTCCCGGGGACCGCTGCGAGCTCCAGGTTCA 600
QY 601 GGGAGGCGTGGCACAGCTGTGATCAAGCGCGCCCGGGGACTCCCATGCTGGGCTGTACAT 660
DB 601 GGGAGGCGTGGCACAGCTGTGATCAAGCGCGCGCCCGGGGACTCCCATGCTGGGCTGTACAT 660
QY 661 GTGGCACCTCTGTGGGACACAGAGAAATCAAGAGTCAAGCTGAGGTTTCAGGTGC 720

Db 661 GTGGCACCTCTGTGGGACACGAGAGAAATAACAGATCAAGTCAAGCTGGAGGTTTCAGGTGC 720
Qy 721 AGAACCCCAAGTCCGCCCCCTGACACTGGGTTCTGGCCCTGTGCGAGCGGTGGTCACTGCTGT 780
Db 721 AGAACCCCAAGTCCGCCCCCTGACACTGGGTTCTGGCCCTGTGCGAGCGGTGGTCACTGCTGT 780
Qy 781 CTTTCATCTCTTGGTGGCTCTGCTCATGTCTGGCTGGTACAGTGGCGCTGTTCCTCAGCA 840
Db 781 CTTTCATCTCTTGGTGGCTCTGCTCATGTCTGGCTGGTACAGTGGCGCTGTTCCTCAGCA 840
Qy 841 AGCCGGGAGAGAGAGTCTTCTCTCTAGAACCCAGATGAAGTGCAGCGCCCTCAGAGC 900
Db 841 AGCCGGGAGAGAGAGTCTTCTCTCTAGAACCCAGATGAAGTGCAGCGCCCTCAGAGC 900
Qy 901 GGGAGCCAGCAGGGGCTGAGCAGAGCCTCCGCTGAATGTGGACCCAGACTCCGAGCC 960
Db 901 GGGAGCCAGCAGGGGCTGAGCAGAGCCTCCGCTGAATGTGGACCCAGACTCCGAGCC 960
Qy 961 CACCCCAAGGCGCTGGACTGTGTCTAAACGCTCACCCTTGAGGCCCTGAGGTGCT 1020
Db 961 CACCCCAAGGCGCTGGACTGTGTCTAAACGCTCACCCTTGAGGCCCTGAGGTGCT 1020
Qy 1021 GTCCCCCCCCAACCTTGTCTCATATGCCGAGAGCCATAGCCGCTGCAAGGCAAGAG 1080
Db 1021 GTCCCCCCCCAACCTTGTCTCATATGCCGAGAGCCATAGCCGCTGCAAGGCAAGAG 1080
Qy 1081 GACACAGAGAGCCAGCCCTGAGTGCAGACCTTGGGTGGCGGGGCTCTCGTCC 1140
Db 1081 GACACAGAGAGCCAGCCCTGAGTGCAGACCTTGGGTGGCGGGGCTCTCGTCC 1140
Qy 1141 CACCCGAGGGGACAGACACCGGCTTGTGGAGGCTGGGCTCTGTGTCAACCACTCC 1200
Db 1141 CACCCGAGGGGACAGACACCGGCTTGTGGAGGCTGGGCTCTGTGTCAACCACTCC 1200
Qy 1201 TGSGGTGCGTGACAGCCCTTCCCTCCACCCCGAGCTTCCCAAGCTCTGCTCTCACT 1260
Db 1201 TGSGGTGCGTGACAGCCCTTCCCTCCACCCCGAGCTTCCCAAGCTCTGCTCTCACT 1260
Qy 1261 TTCCAAAATGGAACCACTCACTCCGAGACACCGGCTTACGAGGACGATGCCCTCC 1320
Db 1261 TTCCAAAATGGAACCACTCACTCCGAGACACCGGCTTACGAGGACGATGCCCTCC 1320
Qy 1321 CTCTGCTCTCATCAACCCAGACACCGGCTTCTTCTGCAACCCAGGCTGGTCCGG 1380
Db 1321 CTCTGCTCTCATCAACCCAGACACCGGCTTCTTCTGCAACCCAGGCTGGTCCGG 1380
Qy 1381 CCCCAGGTGTGGGTGCGCTCTCTCACTCCAGGGCTCCGCGCCCAAGTGAAGGGGCC 1440
Db 1381 CCCCAGGTGTGGGTGCGCTCTCTCACTCCAGGGCTCCGCGCCCAAGTGAAGGGGCC 1440
Qy 1441 CTGCGGAGCCTCAGACACATGGAGTTCAAGGCTGGGGGCTTGGCACAATCCTGTC 1500
Db 1441 CTGCGGAGCCTCAGACACATGGAGTTCAAGGCTGGGGGCTTGGCACAATCCTGTC 1500
Qy 1501 CTTTGGCTATGAGAGGCTTTGGGGCCCTTCGCGGACAGCCCGGGGGCCGAGGTAGGG 1560
Db 1501 CTTTGGCTATGAGAGGCTTTGGGGCCCTTCGCGGACAGCCCGGGGGCCGAGGTAGGG 1560
Qy 1561 TCTGGGGCTTAGAGCTGGGATGCTCTGCTGCGCCACCGCCAGGGGGGCAAGCGAGGCC 1620
Db 1561 TCTGGGGCTTAGAGCTGGGATGCTCTGCTGCGCCACCGCCAGGGGGGCAAGCGAGGCC 1620
Qy 1621 GGGCTGGAGAGGGGGGGGGCTCGGGCTGGGGGCTCAGGTGAAGCTGCTCGGGG 1680
Db 1621 GGGCTGGAGAGGGGGGGGGCTCGGGCTGGGGGCTCAGGTGAAGCTGCTCGGGG 1680
Qy 1681 CTGGTGGCGATCCTCTAGTCCCTCGGCAACCGGGGGTGGTCTGCTGTCGCCACCGCA 1740
Db 1681 CTGGTGGCGATCCTCTAGTCCCTCGGCAACCGGGGGTGGTCTGCTGTCGCCACCGCA 1740
Qy 1741 CTTGCGGAGCCTTTTGGACCCAGATCTGTTTCTGTTTGTCTGCTGCTGCTGCTGCTG 1800
Db 1741 CTTGCGGAGCCTTTTGGACCCAGATCTGTTTCTGTTTGTCTGCTGCTGCTGCTGCTG 1800

Qy 1801 GCCCTTTGATGTCTTTCATCTGTATGGGTGGAATAATCACCGGAATCCCTTTCAGTTTC 1860
Db 1801 GCCCTTTGATGTCTTTCATCTGTATGGGTGGAATAATCACCGGAATCCCTTTCAGTTTC 1860
Qy 1861 TTTGAAAAGTTTCCATGACTCGAATATCTGAATGAAGAAAACAAACCGACTCACAAGCC 1920
Db 1861 TTTGAAAAGTTTCCATGACTCGAATATCTGAATGAAGAAAACAAACCGACTCACAAGCC 1920
Qy 1921 TCCAAAGTAGTCCAAATGCAATTTTAAATGGAATAATCTGAAAGAAACGCTCTT 1980
Db 1921 TCCAAAGTAGTCCAAATGCAATTTTAAATGGAATAATCTGAAAGAAACGCTCTT 1980
Qy 1981 TAGTGCTTTAAGCCCAAAACGTCCTTAAGCGCTCTCGAGATGAAGACGGGGGGAGC 2040
Db 1981 TAGTGCTTTAAGCCCAAAACGTCCTTAAGCGCTCTCGAGATGAAGACGGGGGGAGC 2040
Qy 2041 CCCAGCAGGTGAGACCCCGCAGACCGCGGGCGCCCGGTGACCGAGGCTTCGCACAG 2100
Db 2041 CCCAGCAGGTGAGACCCCGCAGACCGCGGGCGCCCGGTGACCGAGGCTTCGCACAG 2100
Qy 2101 CCGGCGCGCTGAGGCTCGGGCGGAGCCAGGCTTCAAGAGGGCGCGTCTGTCGGG 2160
Db 2101 CCGGCGCGCTGAGGCTCGGGCGGAGCCAGGCTTCAAGAGGGCGCGTCTGTCGGG 2160
Qy 2161 TTTAAATAAGGTTCCGTCGG 2180
Db 2161 TTTAAATAAGGTTCCGTCGG 2180

RESULT 4

US-09-997-165-3
; Sequence 3, Application US/09997165
; Patent No. 6762030
; GENERAL INFORMATION:
; APPLICANT: Lyman, Stewart D.
; TITLE OF INVENTION: LIGAND FOR CD7 AND METHODS OF USE THEREOF
; FILE REFERENCE: 2913-US
; CURRENT APPLICATION NUMBER: US/09/997,165
; PRIORITY FILING DATE: 2001-11-27
; PRIOR APPLICATION NUMBER: PCT/US00/14612
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/136,450
; PRIOR FILING DATE: 1999-05-28
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (119)..(865)
US-09-997-165-3

Query Match 89.5%; Score 1951; DB 4; Length 2000;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1984; Conservative 0; Mismatches 0; Indels 3; Gaps 3;
Qy 196 ATTTTCTTGGGGCTCCGGGGCGGAGGAGCTGCAATCCAGAGGAGCGCTCCAGAGCG 255
Db 1 ATTTTCTTGGGGCTCCGGGGCGGAGGAGCTGCAATCCAGAGGAGCGCTCCAGAGCG 60
Qy 256 GACCCGGGAGTGTTCAGAGCCAGTGAAGGACAGGGGGCCCAAGTCCGACCAAGCAT 315
Db 61 GACCCGGGAGTGTTCAGAGCCAGTGAAGGACAGGGGGCCCAAGTCCGACCAAGCAT 120
Qy 316 GCAGACCTGCCCCCTGGGATTCCTTGGCCACCTTTCCAGGGCCCTTGGAGCCCTCTGTT 375
Db 121 GCAGACCTGCCCCCTGGGATTCCTTGGCCACCTTTCCAGGGCCCTTGGAGCCCTCTGTT 180
Qy 376 TTTGGCTGCTCTCTTGGAGTGTCTCAGATGAAGGCTGGGACAGCCCCCATCTGCACAGAGG 435

181	TTTGCTGCTCTCTGAGTGTCTAGATGAAGGCTGGACAGCCCATCTGCACAGAGG	240	1260	ACACACTGAGTTTCTGAGGCTGGGGGGGCTTTGGCACAATACCTGTCCTTGGCTATGAGCA	1319
436	GGTAGTCTCTGTGTTTGGGGGAGAACACCGTCTATGTCTGTGCAACATCTCCACGCTT	495	1516	GGCTTTGGGGGCTTCCGGCAGCCCGGGGCGGAGGTAGGTCTTGGGGGCTTAGAG	1575
241	GGTAGTCTCTGTGTTTGGGGGAGAACACCGTCTATGTCTGTGCAACATCTCCACGCTT	300	1320	GGCTTTGGGGGCTTCCGGCAGCCCGGGGCGGAGGTAGGTCTTGGGGGCTTAGAG	1379
496	CTCCATGTCAACATCAAGCTCGTGGCCACCGGAGGAGAGCGCATCTTCAATGAGGT	555	1576	GCTGGGATGGCTCTTGGGCCCCACCGCAGGGGGCAAGCGCAGGCTGGAGCGGC	1635
301	CTCCATGTCAACATCAAGCTCGTGGCCACCGGAGGAGAGCGCATCTTCAATGAGGT	360	1380	GCTGGGATGGCTCTTGGGCCCCACCGCAGGGGGCAAGCGCAGGCTGGAGCGGC	1439
556	GGCTCCAGGCTACTTCTCCCGGACCGGTGGCAGTCCAGGTTGAGGAGCGGTGGCACA	615	1636	GGCGGGCTCGGGTGGGGGGTCAAGTGGAGCGCTCCCGGGCTGGTCCGCGCATCCC	1695
361	GGCTCCAGGCTACTTCTCCCGGACCGGTGGCAGTCCAGGTTGAGGAGCGGTGGCACA	420	1440	GGCGGGCTCGGGTGGGGGGTCAAGTGGAGCGCTCCCGGGCTGGTCCGCGCATCCC	1499
616	GCTGGTGAATAAGGGGCGGGGATCCCATGCTGGGCTGTACATGTGGCACTCGTGGG	675	1696	TCACTCCCTCGGCACCGGGGGTCTCCCTCGTGGCCACCGCATCTTGTATGTCCT	1755
421	GCTGGTGAATAAGGGGCGGGGATCCCATGCTGGGCTGTACATGTGGCACTCGTGGG	480	1500	TCACTCCCTCGGCACCGGGGGTCTCCCTCGTGGCCACCGCATCTTGTATGTCCT	1559
676	ACACACAGAGAAATAACAGACAAATCAAGTGGAGGTTTCAAGTGCAGAACCCAGTCCGC	735	1756	TGACACCCAGATCTGTTTCAATGCTTTTGTCTTGTCTGTCACCTGGGGGGGCTTTTGTATGTCCT	1815
481	ACACACAGAGAAATAACAGACAAATCAAGTGGAGGTTTCAAGTGCAGAACCCAGTCCGC	540	1560	TGACACCCAGATCTGTTTCAATGCTTTTGTCTTGTCTGTCACCTGGGGGGGCTTTTGTATGTCCT	1619
736	CCCTGACACTGGGTTCTGGCTGTGCGAGGTTGCTACTGCTGCTTTCATCTCTTGT	795	1816	CATCTGTATGGGTGGGAAATCACCGGAAATCCCTCTTCAAGTCTTGTGAAAGTTCOA	1875
541	CCCTGACACTGGGTTCTGGCTGTGCGAGGTTGCTACTGCTGCTTTCATCTCTTGT	600	1620	CATCTGTATGGGTGGGAAATCACCGGAAATCCCTCTTCAAGTCTTGTGAAAGTTCOA	1679
796	CGCTCTGTCATCTTCCCTGTGACAGTGGCGCTTCCAGCAACCGCGGAGAGAA	855	1876	TGACTCGAATATCTGAAATGAAGAAACAAACCGCATCTCAAAACCTCCAAAGTTCOA	1935
601	CGCTCTGTCATCTTCCCTGTGACAGTGGCGCTTCCAGCAACCGCGGAGAGAA	660	1680	TGACTCGAATATCTGAAATGAAGAAACAAACCGCATCTCAAAACCTCCAAAGTTCOA	1739
856	GTCTCTTCTCTAGAACCCAGATGAAGTTCAGCCCTCAGAGCGGGAGCCAGCAGGG	915	1936	ATGCAATTTTAAATGGAACCAAAATCTGAAAGAAACGTCTTGTAGTGGCTTAAAGCC	1995
661	GTCTCTTCTCTAGAACCCAGATGAAGTTCAGCCCTCAGAGCGGGAGCCAGCAGGG	720	1740	ATGCAATTTTAAATGGAACCAAAATCTGAAAGAAACGTCTTGTAGTGGCTTAAAGCC	1799
916	CCTGAGCAGAGCTCTCGCTGACCTGTGGACCCAGACTCCGAGCCCAAGGCGGCT	975	1996	CCAAAGCTCCTTAAGCGCTCTCGAGATGAAGACGGGGGGAG - CCCCCAGCAGGTGGA	2054
721	CCTGAGCAGAGCTCTCGCTGACCTGTGGACCCAGACTCCGAGCCCAAGGCGGCT	780	1800	CCAAAGCTCCTTAAGCGCTCTCGAGATGAAGACGGGGGGAG - CCCCCAGCAGGTGGA	1859
976	GGCACTGTGTTTAAACCCCTACACATTTGAGGCTTGGAGCTGTGCTCCCTCCACCTT	1035	2055	GACCCCGCAGACCGGGGGGCGCGGTGACCGAGGCTCGCACAGCGCGGCGCTGAG	2114
781	GGCACTGTGTTTAAACCCCTACACATTTGAGGCTTGGAGCTGTGCTCCCTCCACCTT	839	1860	GACCCCGCAGACCGGGGGGCGCGGTGACCGAGGCTCGCACAGCGCGGCGCTGAG	1919
1036	TGTTTCCATATGCCGACACCCATAGCGGCTGCAAGGAGAGAGAGACAGAGAGCCA	1095	2115	GTCGCGGC - GAGCCAGGCTCAAGAGGGGCGCGTGTGTCTCGGTTAAATAAGGTT	2173
840	TGTTTCCATATGCCGACACCCATAGCGGCTGCAAGGAGAGAGAGACAGAGAGCCA	899	1920	GTCGCGGCAGGAGCCAGGCTCGAGAGGGGCGCGTGTGTCTCGGTTAAATAAGGTT	1979
1096	GCCCTGAGTGGACCTTGGGTGGCGGGGCTGGGTCTCTGTCCTCCACCGGAGGACCA	1155	2174	CCGTCGG 2180	
900	GCCCTGAGTGGACCTTGGGTGGCGGGGCTGGGTCTCTGTCCTCCACCGGAGGACCA	959	1980	CCGTCGG 1986	
1156	GACACCGGCTTGTGGAGGCTGGGCTCTGTGTACACCACTCTGGGTGGTGGAGAC	1215			
960	GACACCGGCTTGTGGAGGCTGGGCTCTGTGTACACCACTCTGGGTGGTGGAGAC	1019			
1216	CCTTCCCTCCACCCCGAGGTTTCCAGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	1275			
1020	CCTTCCCTCCACCCCGAGGTTTCCAGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	1079			
1276	ACCTACCTCCGACAGCCGACTTACAGGAGCATGCGCTTCTTCTTCTTCTTCTTCTTCTT	1335			
1080	ACCTACCTCCGACAGCCGACTTACAGGAGCATGCGCTTCTTCTTCTTCTTCTTCTTCTT	1139			
1336	ACCCAGAGCCGAGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	1395			
1140	ACCCAGAGCCGAGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	1199			
1396	CGCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	1455			
1200	CGCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	1259			
1456	ACACACTGAGTTTCTGAGGCTGGGGGGCTTGGCACAATACCTGTCCTTGGCTATGAGCA	1515			

RESULT 5
US-09-949-016-426
; Sequence 426, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 426
; LENGTH: 2000
; TYPE: DNA

; Sequence 2936, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2936
; LENGTH: 2001
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-2936

Query Match 86.7%; Score 1889.6; DB 4; Length 2001;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 1973; Conservative 0; Mismatches 9; Indels 8; Gaps 7;

QY	196	ATTTCTCTGGGCTCGGGGCGGAGAGAGCTGCATCCAGAGGAGCGGTCAGAGCG	255
DB	1	ATTTCTCTGGGCTCGGGGCGGAGAGAGCTGCATCCAGAGGAGCGGTCAGAGCG	60
QY	256	GACCCGGAGTGTTCAGAGCCAGTGACAGACAGAGGAGGAGGAGGAGGAGGAGG	315
DB	61	GACCCGGAGTGTTCAGAGCCAGTGACAGACAGAGGAGGAGGAGGAGGAGGAGG	120
QY	316	GCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	375
DB	121	GCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	180
QY	376	TTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	435
DB	181	TTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	240
QY	436	GTTAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	495
DB	241	GTTAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	300
QY	496	CTCCCATGTCAACATCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	555
DB	301	CTCCCATGTCAACATCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	360
QY	556	GGCTCCAGGCTACTTCTCCGGGAGCGGTCGAGCTCCAGGTTCCAGGAGGCGTGG	615
DB	361	GGCTCCAGGCTACTTCTCCGGGAGCGGTCGAGCTCCAGGTTCCAGGAGGCGTGG	420
QY	616	GCTGGTATCAAGCGCGCGGAGCTCCATGCTGGGCTGTATCATGTGGGACCTCTGG	675
DB	421	GCTGGTATCAAGCGCGCGGAGCTCCATGCTGGGCTGTATCATGTGGGACCTCTGG	480
QY	676	ACACCAGAGAAATACAGCAAGTCAAGCTGGAGGTTTCAGGTTCCAGGAGGCGTGG	735
DB	481	ACACCAGAGAAATACAGCAAGTCAAGCTGGAGGTTTCAGGTTCCAGGAGGCGTGG	540
QY	736	CCCTGACACTGGTCTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGT	795
DB	541	CCCTGACACTGGTCTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGT	600
QY	796	CGCTCTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	855
DB	601	CGCTCTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	660
QY	856	GTTCTCTCTCTAGAACCCCATGAGGTCGACGCTCCAGGCGGAGGCGGAGGCGG	915

DB	661	GTTCCTCTCTAGAACCCCATGAGGTCGACGCTCCAGAGCGGAGGCCAGGAGG	720
QY	916	CTTGACGAGGCTCGGCTGAACTGTGGACCCAGACTCCGAGCCACACCCCAAGCGCT	975
DB	721	CTTGACGAGGCTCGGCTGAACTGTGGACCCAGACTCCGAGCCACACCCCAAGCGCT	780
QY	976	GGCAGCTGTGTTCACAAACCTTCACCACTTGGAGCCCTGGAGCTGTGTCTCC	1035
DB	781	GGCAGCTGTGTTCACAAACCTTCACCACTTGGAGCCCTGGAGCTGTGTCTCC	839
QY	1036	TGTTCCATATGCGCAGACCCATAGCCGCTGCAGGAGGAGGAGGAGGAGGAGG	1095
DB	840	TGTTCCATATGCGCAGACCCATAGCCGCTGCAGGAGGAGGAGGAGGAGGAGG	899
QY	1096	GCCTGAGTGCAGACCTTGGGTGCGGGGCTGGGTCTCTCGTCCACCCGAGGAGCA	1155
DB	900	GCCTGAGTGCAGACCTTGGGTGCGGGGCTGGGTCTCTCGTCCACCCGAGGAGCA	959
QY	1156	GACACCGGCTTGTGTGGCAGGCTGTGTGTACCCACTCTCTGGGTGCGTGAGAC	1215
DB	960	GACACCGGCTTGTGTGGCAGGCTGTGTGTACCCACTCTCTGGGTGCGTGAGAC	1019
QY	1216	CTTCCCTCTCCACCCCGAGGCTTCCAGCTCTCTCTCTCTCTCTCTCTCTCTCT	1275
DB	1020	CTTCCCTCTCCACCCCGAGGCTTCCAGCTCTCTCTCTCTCTCTCTCTCTCTCT	1079
QY	1276	ACCTCACCTCCGAGCACCAGGCTTACAGGAGCGATGCGGCTCTCTCTCTCTCT	1335
DB	1080	ACCTCACCTCCGAGCACCAGGCTTACAGGAGCGATGCGGCTCTCTCTCTCTCT	1139
QY	1336	ACCCACAGACCCGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1395
DB	1140	ACCCACAGACCCGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1199
QY	1396	CGGCT	1455
DB	1200	CGGCT	1259
QY	1456	ACACACTGGAGTTCAGGGC--TGGGGGGGCTTGGGACATACCTCTCTCTCTCTCT	1513
DB	1260	ACACACTGGAGTTCAGGGC--TGGGGGGGCTTGGGACATACCTCTCTCTCTCTCT	1319
QY	1514	CAGGCTTGGGGGCGCTTCCGCGCAGCCCGGGGCGGAGGTAGGGTCTCTCTCTCTCT	1573
DB	1320	CAGGCTTGGGGGCGCTTCCGCGCAGCCCGGGGCGGAGGTAGGGTCTCTCTCTCTCT	1378
QY	1574	AGGCTGGGATGGCT	1633
DB	1379	AGGCTGGGATGGCT	1437
QY	1634	GGGGCGGCGCTCGGGCTGGGGGCTCAGGTGAGCGCTG-CTTCCGGGGGCTGGTCCGCGAT	1692
DB	1438	GGGGCGGCGCTCGGGCTGGGGGCTCAGGTGAGCGCGCCCTCTCCGGGGCTGGACCGCAT	1497
QY	1693	CCCTCAGTCCCTCGGCGACCCCGGGGCTCGCTCTCTCTCTCTCTCTCTCTCTCTCT	1752
DB	1498	CCCTCAGTCCCTCGGCGACCCCGGGGCTCGCTCTCTCTCTCTCTCTCTCTCTCTCT	1557
QY	1753	CTTTGGAACCAAGTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCT	1812
DB	1558	CTTTGGAACCAAGTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCT	1617
QY	1813	CTTTCATCTGATGGGTGGAAAAATCACCGGAAATCCCCCTTTCAGTCTTTTGA	1872
DB	1618	CTTTCATCTGATGGGTGGAAAAATCACCGGAAATCCCCCTTTCAGTCTTTTGA	1677
QY	1873	CCATGACTCGAATATCTGAAATGAGAAACAAACCGACTCACAACCTCCAGTAGCTC	1932
DB	1678	CCATGACTCGAATATCTGAAATGAGAAACAAACCGACTCACAACCTCCAGTAGCTC	1737
QY	1933	CAAAATGCAATTTTAAATGGAACAAATAAATCTGAAAGAAACGCTCTTTAGTGGCTTTAA	1992
DB	1738	CAAAATGCAATTTTAAATGGAACAAATAAATCTGAAAGAAACGCTCTTTAGTGGCTTTAA	1797

QY 1993 GCCCAAAACGTCCTTAAGCGTCTCGAGATGAAGACGGGGGAG-CCCCAGCCAGGT 2051
| | | | |
Db 1798 GCCCAAAACGTCCTTAAGCGTCTCGAGATGAAGACGGGGGAGCCCCAGCCAGGT 1857
| | | | |
QY 2052 GGAGACCCCGCA-GGACGGGGGGCGCCCGGTGACCGAGGCTCCGACGGCGCGCC 2110
| | | | |
Db 1858 GGAGACCCCGAGGACGGGGGGCGCGCGGTGACCGAGGCTCCGACGGCGCGCC 1917
| | | | |
QY 2111 TGAGGGTCGGCGGAGCCAGGCTCAAGAGGGGGCGGTTGTGTCTCGGTTAAATAAG 2170
| | | | |
Db 1918 TGAGGGTCGGCGGAGCCAGGCTCAAGAGGGGGCGGTTGTGTCTCGGTTAAATAAG 1977
| | | | |
QY 2171 GTTCCGTCCG 2180
| | | | |
Db 1978 GTTCCGTCCG 1987
| | | | |

RESULT 7
US-09-949-016-12168
; Sequence 12168, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12168
; LENGTH: 16738
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12168

Query Match 57.0%; Score 1241.6; DB 4; Length 16738;
Best Local Similarity 98.0%; Pred. No. 1.5e-303;
Matches 1331; Conservative 0; Mismatches 19; Indels 8; Gaps 7;
QY 828 GCTGTTCACGCAAGCCGGGAGAGAGTCTTCTCTAGAACCCCGAGATGAAGTGG 887
| | | | |
Db 13370 GCACCTCTCTCCCTCGCAGAAAGTCTTCTCTAGAACCCCGAGATGAAGTGG 13429
| | | | |
QY 888 CAGCCCTCAGAGCGGAGCCAGCAGGCGCTGAGCAGAGCCTCCGCTGAACCTGTGACCC 947
| | | | |
Db 13430 CAGCCCTCAGAGCGGAGCCAGCAGGCGCTGAGCAGAGCCTCCGCTGAACCTGTGACCC 13489
| | | | |
QY 948 CAGACTCCGAGCCCAACCCCAAGCGCGCTGGCACTGGTGTTCAAACCTCACCACCTTGGAG 1007
| | | | |
Db 13490 CAGACTCCGAGCCCAACCCCAAGCGCGCTGGCACTGGTGTTCAAACCTCACCACCTTGGAG 13549
| | | | |
QY 1008 CCCTGGAGCTGTGTCTCCCGCCCAACCTTGTTCATATGCGGAGAGCCCATAGCGGCT 1067
| | | | |
Db 13550 CCCTGGAGCTGTGT-CCCGCCCAACCTTGTTCATATGCGGAGAGCCCATAGCGGCT 13608
| | | | |
QY 1068 GCAAGCGAGAGGACACAGGAGAGCCAGCCCTGAGTCCGACCTTGGTGGCGGGCT 1127
| | | | |
Db 13609 GCAAGCGAGAGGACACAGGAGAGCCAGCCCTGAGTCCGACCTTGGTGGCGGGCT 13668
| | | | |
QY 1128 GGGTCTCTGTCACCCCGGGGACACAGACACCGGCTTGTCTTGGCAGGCTGGGCTCTG 1187
| | | | |
Db 13669 GGGTCTCTGTCACCCCGGGGACACAGACACCGGCTTGTCTTGGCAGGCTGGGCTCTG 13728
| | | | |
QY 1188 TGTACCACTCTCTGGTGGTGGAGACCTTCCCTCACCCCGAGGCTTCCAGCT 1247
| | | | |

Db 13729 TGTACCACTCTCTGGTGGTGGAGACCTTCCCTCCACCCCGAGTCTTCCAGCT 13788
| | | | |
QY 1248 CTGCTTCTCAGTTCCTCAAAATGGAACACACCTCACCTCGCAGCAGCAGGCTTACCAAGA 1307
| | | | |
Db 13789 CTGCTTCTCAGTTCCTCAAAATGGAACACACCTCACCTCGCAGCAGCAGGCTTACCAAGA 13848
| | | | |
QY 1308 CGCATGCCCTTCTCTCTCCCTCATCAAAACCCACAGACCCGAGCTTCTTTCGACACC 1367
| | | | |
Db 13849 CGCATGCCCTTCTCTCTCCCTCATCAAAACCCACAGACCCGAGCTTCTTTCGACACC 13908
| | | | |
QY 1368 CAGGCTGGTCCGGCCCGGAGGTGGGGTCCGCTCTCTCCACCTCCAGGGCTCCGGGCCA 1427
| | | | |
Db 13909 CAGGCTGGTCCGGCCCGGAGGTGGGGTCCGCTCTCTCCACCTCCAGGGCTCCGGGCCA 13968
| | | | |
QY 1428 AGTGAGGGGGCCCTGCGGAGGCTCAGACACACTGGAGTTTCCAGGCT--TGCGGGGGCT 1485
| | | | |
Db 13969 AGTGAGGGGGCCCTGCGGAGGCTCAGACACACTCCAGTTTCCAGGGCTTGGGGGGCTT 14028
| | | | |
QY 1486 TGGCACATACCTGTCTTGGCTATGAGCAGGCTTGGGGGCTTTCGCGGCGACGCCCG 1545
| | | | |
Db 14029 GGCACATACCTGTCTTGGCTATGAGCAGGCTTGGGGGCTTTCGCGGCGACGCCCG 14088
| | | | |
QY 1546 GGGCGGAGGTAGGTCTGGGGGCTTAGAGGCTGGGATGGCTTCCGCGGCCACCGCCAGG 1605
| | | | |
Db 14089 GGGCGGAGGTAGGTCTGGGGGCTTAGAGGCTGGGATGGCTTCCGCGGCCACCGCCAGG 14147
| | | | |
QY 1606 GGGCAAGCGCAGGCGCGGCTGGAGCGCGCGGCTGGGGTGGGGGTTCAGGTGG 1665
| | | | |
Db 14148 GGGC-AGCGCAGGCGCGGCTGGAGCGCGCGGCTGGGGTGGGGGTTCAGGTGG 14206
| | | | |
QY 1666 AGCTG-CTTCGGGGCTGGTGGCGCATCCCTCAGTCCCTCGGCGCACCGGGGGTGGCTC 1724
| | | | |
Db 14207 AGCGCGCTTCCGGGGCTGGAGCGCATCCCTCAGTCCCTCGGCGCACCGGGGGTGGCTC 14266
| | | | |
QY 1725 CCTGTGCGCCACCGCAGCTTGGAGCGCTTGGAGCCAGATCTGTTCATGTCTTGTCT 1784
| | | | |
Db 14267 CCTGTGCGCCACCGCAGCTTGGAGCGCTTGGAGCCAGATCTGTTCATGTCTTGTCT 14326
| | | | |
QY 1785 TCGTCACTGCGCGGGGGCTTGTGATGTCTTCATCTGTATGGGTGGAAAAATCACCGG 1844
| | | | |
Db 14327 TCGTCACTGCGCGGGGGCTTGTGATGTCTTCATCTGTATGGGTGGAAAAATCACCGG 14386
| | | | |
QY 1845 AATCCCTTCTCAGTTCCTTGAAGTTCATGATCGAATATCTGAATATGAAGAAAAACA 1904
| | | | |
Db 14387 AATCCCTTCTCAGTTCCTTGAAGTTCATGATCGAATATCTGAATATGAAGAAAAACA 14446
| | | | |
QY 1905 AACCGACTCACAACTCCAGTCTCAAAATGCAATTTTAAATGGAACAAAAAT 1964
| | | | |
Db 14447 AACCGACTCACAACTCCAGTCTCAAAATGCAATTTTAAATGGAACAAAAAT 14506
| | | | |
QY 1965 CTGAAGAAACGTCCTTGTAGTGGCTTAAAGCCCAAAACGCTCCCTAAGGGCTCTCGAGAT 2024
| | | | |
Db 14507 CTGAAGAAACGTCCTTGTAGTGGCTTAAAGCCCAAAACGCTCCCTAAGGGCTCTCGAGAT 14566
| | | | |
QY 2025 GAAGACGGGGGGAG-CCCCAGCCAGGTGGAGACCCCGCA-GGACCGCGCGCGCCCGT 2082
| | | | |
Db 14567 GAAGACGGGGGGAGCCCGCAGCAGGTGGAGACCCCGCAGCGAGCGCGCGCGCGT 14626
| | | | |
QY 2083 GACCGAGGCTTCGACACAGCGCGCGCTTGGGGTGGGGCGGAGCCAGGCTCAAGAGGG 2142
| | | | |
Db 14627 GACCGAGGCTTCGACACAGCGCGCGCTTGGGGTGGGGCGGAGCCAGGCTCAAGAGGG 14686
| | | | |
QY 2143 GCGGTTTGTCTCGGGTTAAATAAGTTTCCGTCG 2180
| | | | |
Db 14687 GCGGTTTGTCTCGGGTTAAATAAGTTTCCGTCG 14724
| | | | |

RESULT 8
US-09-949-016-14678
; Sequence 14678, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.

14089 GGGGCGAGGTAGGCTC-GGGGGCTTAGAGGCTGGGATGGCTCTTGCGCCCAACCGCCAGG 14147
Qy : 1606 GGGCAAGCCAGCGCGGCTGGAGGCGCGGGCTCGGGCTGGGGGTGAGGTGG 1665
14148 GGGC-AGCGAGGCGGGCTGGAGGCGGGGCTGGGGCTGGGGGTGAGGTGG 14206
Qy 1666 ACGCTG-CCTCCGGGGCTGGTCCGCGATCCCTCAGTCCCTCGGCCACCGGGGGTGGCTC 1724
14207 ACGCGCCCTCCGGGGCTGGAGCGGATCCCTCAGTCCCTCGGCCACCGGGGGTGGCTC 14266
Qy 1725 CTTGTCGCGCCACCGCGACCTCCGAGCTCTTTGGAGCCGAGATCTGTTTCATGCTTTTGTCT 1784
14267 CTTGTCGCGCCACCGCGACCTCCGAGCTCTTTGGAGCCGAGATCTGTTTCATGCTTTTGTCT 14326
Qy 1785 TCGTCACTGCGGGGGCGCTTTGATGCTCTTCAATCTGATGGGGTGGAAAAATCACCGGG 1844
14327 TCGTCACTGCGGGGGCGCTTTGATGCTCTTCAATCTGATGGGGTGGAAAAATCACCGGG 14386
Qy 1845 AATCCCTCTTCAAGTTCTTTGAAAAAGTTCCATGATCGAATATCTGAATGAAGAAACA 1904
14387 AATCCCTCTTCAAGTTCTTTGAAAAAGTTCCATGATCGAATATCTGAATGAAGAAACA 14446
Qy 1905 AACGACTCAAAACCTCCAGTAGCTCCAAATGCAATTTTAAATGGAACCAAAAT 1964
14447 AACGACTCAAAACCTCCAGTAGCTCCAAATGCAATTTTAAATGGAACCAAAAT 14506
Qy 1965 CTGAAGAAACCTCTTTAGTGGCTTTAAGCCCAACCAACGCTCCCTAAGCGTCTCTGAGAT 2024
14507 CTGAAGAAACCTCTTTAGTGGCTTTAAGCCCAACCAACGCTCCCTAAGCGTCTCTGAGAT 14566
Qy 2025 GAAGAGCGGGGGAG-CCCGAGCCAGTGGAGAGCCCGCA-GGACGCGGGGGCGCCGCT 2082
14567 GAAGAGCGGGGGAG-CCCGAGCCAGTGGAGAGCCCGCA-GGACGCGGGGGCGCCGCT 14626
Qy 2083 GACGAGCGCTCGACAGCGCGCCCTGAGGCTCGGGCGGAGCCAGCGGTCCAAGAGGG 2142
14627 GACGAGCGCTCGACAGCGCGCCCTGAGGCTCGGGCGGAGCCAGCGGTCCAAGAGGG 14686
Qy 2143 GCGGCTTTGTCTCGGGTTAAATAAGGTTCCGCTCG 2180
Db : 14687 GCGGCTTTGTCTCGGGTTAAATAAGGTTCCGCTCG 14724

RESULT 9
US-09-949-016-26333/c
; Sequence 26333, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26333
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-26333

Query Match 26.0%; Score 566; DB 4; Length 601;
Best Local Similarity 98.3%; Pred. No. 2,9e-133;
Matches 592; Conservative 1; Mismatches 6; Indels 3; Gaps 2;
Qy 906 CCCAGCAGGCGCTTGGAGAGCGCTCCGCTGGAAGTGTGGAGCCCGAGACTCCGAGCCCAACC 965

14089 GGGGCGAGGTAGGCTC-GGGGGCTTAGAGGCTGGGATGGCTCTTGCGCCCAACCGCCAGG 14147
Qy : 1606 GGGCAAGCCAGCGCGGCTGGAGGCGCGGGCTCGGGCTGGGGGTGAGGTGG 1665
14148 GGGC-AGCGAGGCGGGCTGGAGGCGGGGCTGGGGCTGGGGGTGAGGTGG 14206
Qy 1666 ACGCTG-CCTCCGGGGCTGGTCCGCGATCCCTCAGTCCCTCGGCCACCGGGGGTGGCTC 1724
14207 ACGCGCCCTCCGGGGCTGGAGCGGATCCCTCAGTCCCTCGGCCACCGGGGGTGGCTC 14266
Qy 1725 CTTGTCGCGCCACCGCGACCTCCGAGCTCTTTGGAGCCGAGATCTGTTTCATGCTTTTGTCT 1784
14267 CTTGTCGCGCCACCGCGACCTCCGAGCTCTTTGGAGCCGAGATCTGTTTCATGCTTTTGTCT 14326
Qy 1785 TCGTCACTGCGGGGGCGCTTTGATGCTCTTCAATCTGATGGGGTGGAAAAATCACCGGG 1844
14327 TCGTCACTGCGGGGGCGCTTTGATGCTCTTCAATCTGATGGGGTGGAAAAATCACCGGG 14386
Qy 1845 AATCCCTCTTCAAGTTCTTTGAAAAAGTTCCATGATCGAATATCTGAATGAAGAAACA 1904
14387 AATCCCTCTTCAAGTTCTTTGAAAAAGTTCCATGATCGAATATCTGAATGAAGAAACA 14446
Qy 1905 AACGACTCAAAACCTCCAGTAGCTCCAAATGCAATTTTAAATGGAACCAAAAT 1964
14447 AACGACTCAAAACCTCCAGTAGCTCCAAATGCAATTTTAAATGGAACCAAAAT 14506
Qy 1965 CTGAAGAAACCTCTTTAGTGGCTTTAAGCCCAACCAACGCTCCCTAAGCGTCTCTGAGAT 2024
14507 CTGAAGAAACCTCTTTAGTGGCTTTAAGCCCAACCAACGCTCCCTAAGCGTCTCTGAGAT 14566
Qy 2025 GAAGAGCGGGGGAG-CCCGAGCCAGTGGAGAGCCCGCA-GGACGCGGGGGCGCCGCT 2082
14567 GAAGAGCGGGGGAG-CCCGAGCCAGTGGAGAGCCCGCA-GGACGCGGGGGCGCCGCT 14626
Qy 2083 GACGAGCGCTCGACAGCGCGCCCTGAGGCTCGGGCGGAGCCAGCGGTCCAAGAGGG 2142
14627 GACGAGCGCTCGACAGCGCGCCCTGAGGCTCGGGCGGAGCCAGCGGTCCAAGAGGG 14686
Qy 2143 GCGGCTTTGTCTCGGGTTAAATAAGGTTCCGCTCG 2180
Db : 14687 GCGGCTTTGTCTCGGGTTAAATAAGGTTCCGCTCG 14724

RESULT 9
US-09-949-016-26333/c
; Sequence 26333, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26333
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-26333

Query Match 26.0%; Score 566; DB 4; Length 601;
Best Local Similarity 98.3%; Pred. No. 2,9e-133;
Matches 592; Conservative 1; Mismatches 6; Indels 3; Gaps 2;
Qy 906 CCCAGCAGGCGCTTGGAGAGCGCTCCGCTGGAAGTGTGGAGCCCGAGACTCCGAGCCCAACC 965

Db 601 CCAGAGGGCTGAGCAGAGCCTCCGCTGAACCTGTGACCCAGACCTCCGAGCCACCC 542
Qy 966 CAAGGCCGCTGGCACTGGTGTCAAAACCTTCAACCTTGGAGCCCTGGAGCTGTGCTCC 1025
Db 541 CAAGGCCGCTGGCACTGGTGTCAAAACCTTCAACCTTGGAGCCCTGGAGCTGTGCTCC 483
Qy 1026 CCCCACCCCTTTTTCATATATGCGCAGACCCATAGCCGCTTGAAGGAGAGAGACAC 1085
Db 482 CCCCACCCCTTTTTCATATATGCGCAGACCCATAGCCGCTTGAAGGAGAGAGACAC 423
Qy 1086 AGGAGAGCAGCCCTGAGTGGCGGCTTGGGTGGGGGCTGGGTCTCTGCTCCACCC 1145
Db 422 AGGAGAGCAGCCCTGAGTGGCGGCTTGGGTGGGGGCTGGGTCTCTGCTCCACCC 363
Qy 1146 GGAGGGCAGACACCCGCTTGGTGGCAGGCTGGGGCTCTGTGTCAACCACTCCCTGGGT 1205
Db 362 GGAGGGCAGACACCCGCTTGGTGGCAGGCTGGGGCTCTGTGTCAACCACTCCCTGGGT 303
Qy 1206 GGTGCGAGACCTTCCCTTCCACCCCGCAGGTCTTCCAGCTCTGCTTCCCTCAGTTTCCA 1265
Db 302 GYGTGCGAGACCTTCCCTTCCACCCCGCAGGTCTTCCAGCTCTGCTTCCCTCAGTTTCCA 243
Qy 1266 AAATGGAACCACTCAGCTCCGAGACCCGAGCTTACAGGACGATGCCCTCCCTCTG 1325
Db 242 AAATGGAACCACTCAGCTCCGAGACCCGAGCTTACAGGACGATGCCCTCCCTCTG 183
Qy 1326 CCTCATCAAAACCCAGACCCGAGCTCCCTTCTGCGACCCAGGCTGGTCCGGCCCCA 1385
Db 182 CCTCATCAAAACCCAGACCCGAGCTCCCTTCTGCGACCCAGGCTGGTCCGGCCCCA 123
Qy 1386 GGTGCGGCTCGCTCTCTCCACTCCAGGGCTCCGCGCCCAAGTGAGGGGCCCCCTGCC 1445
Db 122 GGTGCGGCTCGCTCTCTCCACTCCAGGGCTCCGCGCCCAAGTGAGGGGCCCCCTGCC 63
Qy 1446 GGAGCTCAGACACACTGAGTTTCAAGGC--TGGGGGGGCTTGGCACAATCTGCTCCCT 1503
Db 62 GGAGCTCAGACACACTGAGTTTCAAGGCCTGTGGGGGGCTTGGGCCACATACCTGTCCCT 3
Qy 1504 TG 1505
Db 2 TG 1

RESULT 10
US-09-949-016-105202/c
; Sequence 105202, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 105202
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-105202

Query Match 26.0%; Score 566; DB 4; Length 601;
Best Local Similarity 98.3%; Pred. No. 2.9e-133; Mismatches 6; Indels 3; Gaps 2;
Matches 592; Conservative 1; Mismatches 6; Indels 3; Gaps 2;

Qy 906 CCAGAGGGCTGAGCAGAGCCTCCGCTGAACCTGTGACCCAGACCTCCGAGCCACCC 965
Db 601 CCAGAGGGCTGAGCAGAGCCTCCGCTGAACCTGTGACCCAGACCTCCGAGCCACCC 542
Qy 966 CAAGGCCGCTGGCACTGGTGTCAAAACCTTCAACCTTGGAGCCCTGGAGCTGTGCTCC 1025
Db 541 CAAGGCCGCTGGCACTGGTGTCAAAACCTTCAACCTTGGAGCCCTGGAGCTGTGCTCC 483
Qy 1026 CCCCACCCCTTTTTCATATATGCGCAGACCCATAGCCGCTTGAAGGAGAGAGACAC 1085
Db 482 CCCCACCCCTTTTTCATATATGCGCAGACCCATAGCCGCTTGAAGGAGAGAGACAC 423
Qy 1086 AGGAGAGCAGCCCTGAGTGGCGGCTTGGGTGGGGGCTGGGTCTCTGCTCCACCC 1145
Db 422 AGGAGAGCAGCCCTGAGTGGCGGCTTGGGTGGGGGCTGGGTCTCTGCTCCACCC 363
Qy 1146 GGAGGGCAGACACCCGCTTGGTGGCAGGCTGGGGCTCTGTGTCAACCACTCCCTGGGT 1205
Db 362 GGAGGGCAGACACCCGCTTGGTGGCAGGCTGGGGCTCTGTGTCAACCACTCCCTGGGT 303
Qy 1206 GGTGCGAGACCTTCCCTTCCACCCCGCAGGTCTTCCAGCTCTGCTTCCCTCAGTTTCCA 1265
Db 302 GYGTGCGAGACCTTCCCTTCCACCCCGCAGGTCTTCCAGCTCTGCTTCCCTCAGTTTCCA 243
Qy 1266 AAATGGAACCACTCAGCTCCGAGACCCGAGCTTACAGGACGATGCCCTCCCTCTG 1325
Db 242 AAATGGAACCACTCAGCTCCGAGACCCGAGCTTACAGGACGATGCCCTCCCTCTG 183
Qy 1326 CCTCATCAAAACCCAGACCCGAGCTCCCTTCTGCGACCCAGGCTGGTCCGGCCCCA 1385
Db 182 CCTCATCAAAACCCAGACCCGAGCTCCCTTCTGCGACCCAGGCTGGTCCGGCCCCA 123
Qy 1386 GGTGCGGCTCGCTCTCTCCACTCCAGGGCTCCGCGCCCAAGTGAGGGGCCCCCTGCC 1445
Db 122 GGTGCGGCTCGCTCTCTCCACTCCAGGGCTCCGCGCCCAAGTGAGGGGCCCCCTGCC 63
Qy 1446 GGAGCTCAGACACACTGAGTTTCAAGGC--TGGGGGGGCTTGGCACAATCTGCTCCCT 1503
Db 62 GGAGCTCAGACACACTGAGTTTCAAGGCCTGTGGGGGGCTTGGGCCACATACCTGTCCCT 3
Qy 1504 TG 1505
Db 2 TG 1

RESULT 11
US-09-949-016-26334/c
; Sequence 26334, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26334
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-26334

Query Match 24.8%; Score 541.4; DB 4; Length 601;
Best Local Similarity 97.8%; Pred. No. 4.8e-127; Mismatches 7; Indels 5; Gaps 4;
Matches 590; Conservative 1; Mismatches 7; Indels 5; Gaps 4;

		Matches 590; Conservative 1; Mismatches 7; Indels 5; Gaps 4;	
QY	1370	GGCTGGTCCGGCCCGAGGTGGGGTCCGCTCTCTCCACTCCAGGGCTCCGCGCCCAAG	1429
	601	GGCTGGTCCGGCCCGAGGTGGGGTCCGCTCTCTCCACTCCAGGGCTCCGCGCCCAAG	542
Db	1430	TGAGGGGGCCCTGCGGAGCCTCAGACACTGGAGTTCAAGGC--TGGGGGGGCCTTG	1487
	541	TGAGGGGGCCCTGCGGAGCCTCAGACACTGGAGTTCAAGGC--TGGGGGGGCCTTG	482
QY	1488	GCACATACCTGTCCTTGGCTATGACAGGCTTTGGGGGCCCTTCCGCGGAGCCCGGG	1547
	481	GCACATACCTGTCCTTGGCTATGACAGGCTTTGGGGGCCCTTCCGCGGAGCCCGGG	422
Db	1548	GGCGGAGGTAGGGTCTGGGGGCTTAGAGGCTGGGATGGCTCTTGGGCCCAACCGCCAGGG	1607
	421	GGCGGAGGTAGGGTCTGGGGGCTTAGAGGCTGGGATGGCTCTTGGGCCCAACCGCCAGGG	363
QY	1608	GCAGCGCAGCGGGGCTGGGAGCGCGCGCGGCTCGGGCTGGGGGTGAGGTGGAC	1667
	362	GC-AGCGCAGCGGGGCTGGGAGCGCGCGCGGCTCGGGCTGGGGGTGAGGTGGAC	304
QY	1668	GCTG-CCTCGGGGCTGGTCCGCGCATCCCTCAGTCCCTCGGCCACCCCGGGGTGCTCCC	1726
	303	GCTG-CCTCGGGGCTGGTCCGCGCATCCCTCAGTCCCTCGGCCACCCCGGGGTGCTCCC	244
Db	1727	TCGTGCCACCGCAGCTGGCGAGCCTCTTTGGAGCCAGATCTGTTCAATGCTTTGCTTC	1786
	243	TCGTGCCACCGCAGCTGGCGAGCCTCTTTGGAGCCAGATCTGTTCAATGCTTTGCTTC	184
QY	1787	GTCACTGGCGGGGCGCTTGGATGCTTCACTGATGGGTGGAATAATCACCGGAA	1846
	183	GTCACTGGCGGGGCGCTTGGATGCTTCACTGATGGGTGGAATAATCACCGGAA	124
QY	1847	TCCCTCTCAGTTCTTTGAAAAAGTTCATGACTCGAATATCTGAAATGAAGAAACAAA	1906
	123	TCCCTCTCAGTTCTTTGAAAAAGTTCATGACTCGAATATCTGAAATGAAGAAACAAA	64
Db	1907	CGACTCAGAACTCCAGTACTCGAATGCAATTTTAAATGAAACAAAAATCT	1966
	63	CGACTCAGAACTCCAGTACTCGAATGCAATTTTAAATGAAACAAAAATCT	4
QY	1967	GAA 1969	
	3	GAA 1	
RESULT 12			
US-09-949-016-105203/c			
; Sequence 105203, Application US/09949016			
; Patent No. 6812339			
; GENERAL INFORMATION:			
; APPLICANT: VENTER, J. Craig et al.			
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED			
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF			
; FILE REFERENCE: CL001307			
; CURRENT APPLICATION NUMBER: US/09/949,016			
; CURRENT FILING DATE: 2000-04-14			
; PRIOR APPLICATION NUMBER: 60/241,755			
; PRIOR FILING DATE: 2000-10-20			
; PRIOR APPLICATION NUMBER: 60/237,768			
; PRIOR FILING DATE: 2000-10-03			
; PRIOR APPLICATION NUMBER: 60/231,498			
; PRIOR FILING DATE: 2000-09-08			
; NUMBER OF SEQ ID NOS: 207012			
; SOFTWARE: FastSeq For Windows Version 4.0			
; SEQ ID NO 105203			
; LENGTH: 601			
; TYPE: DNA			
; ORGANISM: Human			
US-09-949-016-105203			
Query Match 24.8%; Score 541.4; DB 4; Length 601;			
Best Local Similarity 97.8%; Pred. No. 4.8e-127;			

Query Match 7.1%; Score 155.4; DB 4; Length 1112;
Best Local Similarity 60.6%; Pred. No. 3e-29;
Matches 255; Conservative 0; Mismatches 166; Indels 0; Gaps 0;

QY 310 AGCCATGACAGCTGCCCTCCCTGCAATTCCTGGCCACGTTTCCAGGCCCTTTGGGACCT 369
DB 131 AGCCATGCTGGCTACTCTGTAAACATCTCTGGCTGTTTCCAGATGCTCTGGGCT 190

QY 370 CTGTTTTTGGCTGCTCTCTGAGTGTCTAGATGAAGCTGGGACAGCCCATCTGCAC 429
DB 191 CCTTCTACTGGGGCTCCCTGAATGCCCATAACGATGCTCTGGGACGAACCTTTGTGCAC 250

QY 430 AGAGGGGTAGTCTCTGTGTCTTGGGGCGAGAACCCGTCATGTCTCTGCAACATCTCAA 489
DB 251 TGAGCATGAAGTATCTGTAACAGAGGCGCGCTGTGGTGTGATGGCTGTAAATATCTCAA 310

QY 490 CGCCTTCTCCCATGTCAACATCAAGCTGGTCCCGGACGAGGAGGCCATCTTCAA 549
DB 311 CAATCTCAGAGAGCTACCAATGAGTTGGTTACAGTGAAGACTAGCATCTTCAA 370

QY 550 TGAGTGGCTCCAGGCTACTTCTCCCGGACGGCTGGCAGCTCCAGGTTCCAGGGAGGGT 609
DB 371 TCATAGGCTCCAGGAACACTACTTAAGGATTCATGGCAGCTTCATATTCAGAGTCCA 430

QY 610 GGCACAGCTGTGATCAAGGGCGCCGGGACTCCCATGCTGGGCTGTACATGTGGCACC 669
DB 431 GGGCCAGCTGTGTATCACAGATGCTCAGGACAAACACTCAGGGAACACTACTCTGGAAGCT 490

QY 670 CTGGGACACACAGAGAAATAACAGACAAGTCAAGCTGGAGGTTTCAGGTGACAGACCCA 729
DB 491 GCATGGATTCAGGAGAGTTCAAAACCTTCAACCTGACTGTTAATGCGGACAGACAGA 550

QY 730 G 730
DB 551 G 551

RESULT 14
US-09-949-016-26332/c
; Sequence 26332, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26332
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-26332

Query Match 6.4%; Score 138.6; DB 4; Length 601;
Best Local Similarity 99.3%; Pred. No. 4e-25;
Matches 138; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 712 TTCAGGTGCAGAACCCCGAGTCGCCCTTGACACTGGGTTCTGGCTGTGCCAGCGGTGT 771
DB 328 TTCAGGTGCAGAACCCCGAGTCGCCCTTGACACTGGGTTCTGGCTGTGCCAGCGGTGT 269

QY 772 CACTGCTGTCTTCATCTTGGTCTGGTCTGTCTATGTTTCGCTGTGTACAGGTGCCGCTG 831
DB 268 CACTGCTGTCTTCATCTTGGTCTGGTCTGTCTATGTTTCGCTGTGTACAGGTGCCGCTG 209

QY 832 TTCCAGCAACGCCCGGAG 850
DB 208 TTCCAGCAACGCCCGGAG 190

RESULT 15

US-09-949-016-105201/c
; Sequence 105201, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 105201
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-105201

Query Match 6.4%; Score 138.6; DB 4; Length 601;
Best Local Similarity 99.3%; Pred. No. 4e-25;
Matches 138; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 712 TTCAGGTGCAGAACCCCGAGTCGCCCTTGACACTGGGTTCTGGCTGTGCCAGCGGTGT 771
DB 328 TTCAGGTGCAGAACCCCGAGTCGCCCTTGACACTGGGTTCTGGCTGTGCCAGCGGTGT 269

QY 772 CACTGCTGTCTTCATCTTGGTCTGGTCTGTCTATGTTTCGCTGTGTACAGGTGCCGCTG 831
DB 268 CACTGCTGTCTTCATCTTGGTCTGGTCTGTCTATGTTTCGCTGTGTACAGGTGCCGCTG 209

QY 832 TTCCAGCAACGCCCGGAG 850
DB 208 TTCCAGCAACGCCCGGAG 190

Search completed: March 20, 2005, 06:56:11
Job time : 386 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 20, 2005, 04:49:49 ; Search time 1256 Seconds
(without alignments)
10332.824 Million cell updates/sec

Title: US-10-080-522-2

Perfect score: 2180

Sequence: 1 ATTCTCGTCTCTTAGCGT.....TTAAATAAGTTCGTCG 2180

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 554816 seqs, 2976611598 residues

Total number of hits satisfying chosen parameters: 11089632

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
- 18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
- 19: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 20: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
- 21: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 22: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2180	100.0	2180	14	US-10-080-522-2
2	1951	89.5	2000	9	US-09-954-456-1237
3	1951	89.5	2000	9	US-09-997-165-3
4	1951	89.5	2000	18	US-10-775-169-97
5	912.8	41.9	1064	9	US-09-799-777-109
c 6	222	10.2	3026	18	US-10-676-248B-245
7	155.4	7.1	1112	9	US-09-997-165-7
8	60	2.8	60	10	US-09-908-975-11235
9	54	2.5	100301	18	US-10-450-826-83
10	54	2.5	100301	18	US-10-723-860-3574
c 11	52	2.4	606	17	US-10-424-599-95310

c 12	51.2	2.3	5452	15	US-10-017-161-1481	Sequence 1481, Ap
c 13	51.2	2.3	5452	17	US-10-292-798-1189	Sequence 1189, Ap
c 14	50.4	2.3	805	17	US-10-424-599-61903	Sequence 61903, A
c 15	50.2	2.3	671	14	US-10-184-644-346	Sequence 346, App
c 16	50.2	2.3	671	14	US-10-184-644-346	Sequence 346, App
c 17	50	2.3	728	18	US-10-425-115-183861	Sequence 183861, A
c 18	49.6	2.3	645	18	US-10-767-701-3807	Sequence 3807, Ap
c 19	49.6	2.3	763	18	US-10-437-963-49370	Sequence 49370, A
c 20	49.6	2.3	882	14	US-10-184-644-574	Sequence 574, App
c 21	49.6	2.3	882	14	US-10-184-644-574	Sequence 574, App
c 22	49.4	2.3	4667	18	US-10-723-860-5759	Sequence 5759, Ap
c 23	49.4	2.3	5452	15	US-10-017-161-1481	Sequence 1481, Ap
c 24	49.4	2.3	5452	17	US-10-292-798-1189	Sequence 1189, Ap
c 25	49.2	2.3	581	18	US-10-425-115-93481	Sequence 93481, A
c 26	49.2	2.3	581	18	US-10-425-115-93481	Sequence 93481, A
c 27	49.2	2.3	1506	18	US-10-437-963-38913	Sequence 38913, A
c 28	49	2.2	805	17	US-10-424-599-61903	Sequence 61903, A
c 29	48.8	2.2	959	18	US-10-437-963-48694	Sequence 48694, A
c 30	48.8	2.2	1533	18	US-10-739-930-4713	Sequence 4713, Ap
c 31	48.6	2.2	806	18	US-10-425-115-123827	Sequence 123827, A
c 32	48.6	2.2	972	18	US-10-425-115-17653	Sequence 17653, A
c 33	48.4	2.2	486	18	US-10-425-115-181872	Sequence 181872, A
c 34	48.4	2.2	737	18	US-10-425-115-65924	Sequence 65924, A
c 35	48.4	2.2	3897	17	US-10-213-796-20	Sequence 20, Appl
c 36	48.4	2.2	22210	17	US-10-213-796-19	Sequence 19, Appl
c 37	48.2	2.2	594	14	US-10-123-155-10	Sequence 10, Appl
c 38	48.2	2.2	594	15	US-10-146-731-10	Sequence 10, Appl
c 39	48.2	2.2	594	15	US-10-140-472-10	Sequence 10, Appl
c 40	48.2	2.2	594	15	US-10-141-761-10	Sequence 10, Appl
c 41	48.2	2.2	594	16	US-10-142-885-10	Sequence 10, Appl
c 42	48.2	2.2	594	16	US-10-158-790-10	Sequence 10, Appl
c 43	48.2	2.2	594	17	US-10-137-871-10	Sequence 10, Appl
c 44	48.2	2.2	594	17	US-10-140-923-10	Sequence 10, Appl
c 45	48.2	2.2	594	17	US-10-141-756-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1

US-10-080-522-2
; Sequence 2, Application US/10080522
; Publication No. US20030096326A1
; GENERAL INFORMATION:
; APPLICANT: KAUFMAN, RUSSEL E.
; SLENTZ-KESLER, KIMBERLY
; TITLE OF INVENTION: GENE PRODUCT OVER EXPRESSED IN CANCER CELLS
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/080,522
; FILING DATE: 25-Feb-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/539,774
; FILING DATE: 31-MAR-2000
; APPLICATION NUMBER: US 09/210,474
; FILING DATE: 14-DEC-1998
; APPLICATION NUMBER: US 08/755,559
; FILING DATE: 22-NOV-1996
; ATTORNEY/AGENT INFORMATION:

NAME: WILSON, MARY J.	
REGISTRATION NUMBER: 32,955	
REFERENCE/DOCKET NUMBER: 1579-645	
TELECOMMUNICATION INFORMATION:	
TELEPHONE: (703) 816-4000	
TELEFAX: (703) 816-4100	
INFORMATION FOR SEQ ID NO: 2:	
SEQUENCE CHARACTERISTICS:	
LENGTH: 2180 base pairs	
TYPE: nucleic acid	
STRANDEDNESS: single	
TOPOLOGY: linear	
MOLECULE TYPE: DNA (genomic)	
SEQUENCE DESCRIPTION: SEQ ID NO: 2:	
US-10-080-522-2	
Query Match 100.0%; Score 2180; DB 14; Length 2180;	
Best Local Similarity 100.0%; Pred. No. 0;	
Matches 2180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Y	1 ATTCTGCTTCTTTAGCGTGAACGCGGTGCGGTGCGTCCCGTGAATAATAAATTCAC 60
o	1 ATTCTGCTTCTTTAGCGTGAACGCGGTGCGGTGCGTCCCGTGAATAATAAATTCAC 60
Y	61 CGTCACGCTTGTGTGAACGCGGGTGGTTCOCGAAACTTGGAGGCTTCCCGTAAACCCAG 120
o	61 CGTCACGCTTGTGTGAACGCGGGTGGTTCOCGAAACTTGGAGGCTTCCCGTAAACCCAG 120
Y	121 CTCCTTCTCATCTGGAGGTGGTCCCGCGGGTCCGCGGCTCTCTCCCTGGGCCCTC 180
o	121 CTCCTTCTCATCTGGAGGTGGTCCCGCGGGTCCGCGGCTCTCTCCCTGGGCCCTC 180
Y	181 CCTCTGCTGCTTTCAATTTCTTGGGGTCCGGGCGCGGAGAGCTGCATCCACAGAGA 240
o	181 CCTCTGCTGCTTTCAATTTCTTGGGGTCCGGGCGCGGAGAGTGCATCCACAGAGA 240
Y	241 GCGGTCACAGAGCCGAGACCCGGAGTGTTCAGAGCCAGTGAAGAGCCAGAGGCGGCCA 300
o	241 GCGGTCACAGAGCCGAGACCCGGAGTGTTCAGAGCCAGTGAAGAGCCAGAGGCGGCCA 300
Y	301 AGTCCCAACAGCCATGACAGCCTGCCCTTGGCAATTCCTTGGCCAGGTTTCCAGGCCCT 360
o	301 AGTCCCAACAGCCATGACAGCCTGCCCTTGGCAATTCCTTGGCCAGGTTTCCAGGCCCT 360
Y	361 TGGGACCTCTCTTTTGGTGGCTGCTTGGTGTCTCAGAGTGAAGGCTGGGACAGCCC 420
o	361 TGGGACCTCTCTTTTGGTGGCTGCTTGGTGTCTCAGAGTGAAGGCTGGGACAGCCC 420
Y	421 CATCTGCACAGAGGGGTAGTCTCTGTGCTTGGGGCGAGAAACACCGTCAATGCTCTGCAA 480
o	421 CATCTGCACAGAGGGGTAGTCTCTGTGCTTGGGGCGAGAAACACCGTCAATGCTCTGCAA 480
Y	481 CATCTCAACAGCCTTCTCCATGTCAACATCAAGCTCGTGGCCGACGCGGAGAGAGCGC 540
o	481 CATCTCAACAGCCTTCTCCATGTCAACATCAAGCTCGTGGCCGACGCGGAGAGAGCGC 540
Y	541 CATCTCAATAGTGGTGGCTCAGGCTACTTCTCCCGGACGCGTGGCAGTCCAGGTTCA 600
o	541 CATCTCAATAGTGGTGGCTCAGGCTACTTCTCCCGGACGCGTGGCAGTCCAGGTTCA 600
Y	601 GGGAGGCGTGCGCAGCTGTGTATCAAAAGCGCCCGGACTCCCATGCTGGGCTGTACAT 660
o	601 GGGAGGCGTGCGCAGCTGTGTATCAAAAGCGCCCGGACTCCCATGCTGGGCTGTACAT 660
Y	661 GTGGCACTGTGGGACACAGAGAAATAACAGACAAGTCAAGCTGAGGTTTCAGGTGC 720
o	661 GTGGCACTGTGGGACACAGAGAAATAACAGACAAGTCAAGCTGAGGTTTCAGGTGC 720
Y	721 AGAACCCAGTCCGCCCTGACACTGGGTTCTGGCTGTGGCCAGCGGTGTGCTCACTGCTGT 780
o	721 AGAACCCAGTCCGCCCTGACACTGGGTTCTGGCTGTGGCCAGCGGTGTGCTCACTGCTGT 780
Y	781 CTTTCATCTTGTGCTCTGGTCAATGTTCCGCTGGTACAGGTGCGGCTGTTCCACAGCA 840

Db 1861 TTGTGAAAGTTCCATGACTGAATATCTGAATGAAGAAAACAAACCGACTCACAAACC 1920
Qy 1921 TCCAGTAGCTCAAAATGCAATTTTTTAAATGGAACAAAAATCTGAAGAAACGTCTTT 1980
Db 1921 TCCAGTAGCTCAAAATGCAATTTTTTAAATGGAACAAAAATCTGAAGAAACGTCTTT 1980
Qy 1981 TAGTGGCTTTAAGCCCAAAACGTCCCTTAAGGGCTCTCGAGATGAAGACGGGGGGAGC 2040
Db 1981 TAGTGGCTTTAAGCCCAAAACGTCCCTTAAGGGCTCTCGAGATGAAGACGGGGGGAGC 2040
Qy 2041 CCAGCCAGGTGAGACCCCGCAGACGCGGGCGCGCGCGCGGTGACCGAGGCTTCGCACAG 2100
Db 2041 CCAGCCAGGTGAGACCCCGCAGACGCGGGCGCGCGCGGTGACCGAGGCTTCGCACAG 2100
Qy 2101 CCGGGCGCCCTGAGGGTCCGGGCGGACGACGAGGTCCAAAGAGGGCGGCTTGTGTCTCGGG 2160
Db 2101 CCGGGCGCCCTGAGGGTCCGGGCGGACGAGGTCCAAAGAGGGCGGCTTGTGTCTCGGG 2160
Qy 2161 TTAAATAAAGTTCCGTCCG 2180
Db 2161 TTAAATAAAGTTCCGTCCG 2180

RESULT 2
US-09-954-456-1237
; Sequence 1237, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cand
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; PRIORITY FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1237
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-1237

Query Match 89.5%; Score 1951; DB 9; Length 2000;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1984; Conservative 0; Mismatches 0; Indels 3; Gaps 3;

Qy 196 ATTTTCTCGGGCTCCGGGCGGAGAAAGCTGCATCCAGAGGAGCGCGTCCAGAGCG 255
Db 1 ATTTTCTCGGGCTCCGGGCGGAGAAAGCTGCATCCAGAGGAGCGCGTCCAGAGCG 60
Qy 256 GACCCGGGAGTGTTCAGAGCCAGTGAAGGACCGAGGCGCCCAAGTCCGACCCACCAT 315
Db 61 GACCCGGGAGTGTTCAGAGCCAGTGAAGGACCGAGGCGCCCAAGTCCGACCCACCAT 120

Qy 316 GCAGACCTGCCCCCTTGGCAATTCCTGGCCACGTTTCCAGGCCCCCTTGGACCCCTCTGT 375
Db 121 GCAGACCTGCCCCCTTGGCAATTCCTGGCCACGTTTCCAGGCCCCCTTGGACCCCTCTGT 180
Qy 376 TTTGGCTGCTCTTGTAGTGTCTAGATGAAGGCTGGGACGCCCATCTGACAGAGGG 435
Db 181 TTTGGCTGCTCTTGTAGTGTCTAGATGAAGGCTGGGACGCCCATCTGACAGAGGG 240
Qy 436 GGTAGTCTCTGTGTCTTGGGGCGAACAACCGTCAATGCTCTGCAACATCTCCAAAGCCCTT 495
Db 241 GGTAGTCTCTGTGTCTTGGGGCGAACAACCGTCAATGCTCTGCAACATCTCCAAAGCCCTT 300
Qy 496 TTCCCATGTCAACATCAAGCTGCTGCCCCACGGGACGAGAGCGCCATCTTCAATAGGT 555
Db 301 TTCCCATGTCAACATCAAGCTGCTGCCCCACGGGACGAGAGCGCCATCTTCAATAGGT 360
Qy 556 GGTCCAGGCTACTTCTCCGGGACGGCTGCGAGTCTCCAGTTCAGGAGGGGTGGCACA 615
Db 361 GGTCCAGGCTACTTCTCCGGGACGGCTGCGAGTTCAGGAGGGGTGGCACA 420
Qy 616 GCTGGTGATCAAAAGGCGCCGGGACTCCCATGCTGGGCTGTACATGTGGCAGCTCTGGG 675
Db 421 GCTGGTGATCAAAAGGCGCCGGGACTCCCATGCTGGGCTGTACATGTGGCAGCTCTGGG 480
Qy 676 ACACAGAGAAATAACAGACAAAGTCAAGTCAAGTGGAGGTTTCAAGTGCAGAACCCCAAGTCCGC 735
Db 481 ACACAGAGAAATAACAGACAAAGTCAAGTCAAGTGGAGGTTTCAAGTGCAGAACCCCAAGTCCGC 540
Qy 736 CCTGACACTGGGTTCTGGGCTGTCGAGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 795
Db 541 CCTGACACTGGGTTCTGGGCTGTCGAGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
Qy 796 CGCTCTGGTCAATGTTCCGCTGCTGTCAGAGTCCGCTGTTCCAGCAACCGCGGAGAGAA 855
Db 601 CGCTCTGGTCAATGTTCCGCTGCTGTCAGAGTCCGCTGTTCCAGCAACCGCGGAGAGAA 860
Qy 856 GTTCTTCTCTCTAGAACCCAGATGAAGTCCAGCCCTCAGAGCGGAGCCAGCAGGG 915
Db 661 GTTCTTCTCTCTAGAACCCAGATGAAGTCCAGCCCTCAGAGCGGAGCCAGCAGGG 720
Qy 916 CCTGAGCAGAGGCTCCGCTGAACTGTGGACCCAGACTCCGAGCCCAACCCCAAGGCCGT 975
Db 721 CCTGAGCAGAGGCTCCGCTGAACTGTGGACCCAGACTCCGAGCCCAACCCCAAGGCCGT 780
Qy 976 GGCAGTGTGTTCAAAACCTCACCCTTGGAGCCCTCGAGCTGCTGCTGCTGCTGCTGCT 1035
Db 781 GGCAGTGTGTTCAAAACCTCACCCTTGGAGCCCTCGAGCTGCTGCTGCTGCTGCTGCT 839
Qy 1036 TGTTCCTATATGCGCAGACCCATAGCCGCTTCAAGGCGAGAGGACACAGGAGGCCA 1095
Db 840 TGTTCCTATATGCGCAGACCCATAGCCGCTTCAAGGCGAGAGGACACACAGGAGGCCA 899
Qy 1096 GGCCTGAGTGGGACCTTGGGTGGGGGGGCTGGGTCTCTGCTCCACCCCGAGGGGACA 1155
Db 900 GGCCTGAGTGGGACCTTGGGTGGGGGGGCTGGGTCTCTGCTCCACCCCGAGGGGACA 959
Qy 1156 GACACCGGCTTGTGGCAGGCTGGGCTCTGCTGTCAACCACTCTCTGGGTGGGTGAGAC 1215
Db 960 GACACCGGCTTGTGGCAGGCTGGGCTCTGCTGTCAACCACTCTCTGGGTGGGTGAGAC 1019
Qy 1216 CTTTCCCTTCAACCCCGAGGCTTCTTCAAGTCTGCTTCTCAGTTCCTCAAAATGGAACC 1275
Db 1020 CTTTCCCTTCAACCCCGAGGCTTCTTCAAGTCTGCTTCTCAGTTCCTCAAAATGGAACC 1079
Qy 1276 ACCTCACTCCGAGACCCGAGCTTACAGAGGAGCATGCCCCCTCTGCTGCTGCTGCTGCT 1335
Db 1080 ACCTCACTCCGAGACCCGAGCTTACAGAGGAGCATGCCCCCTCTGCTGCTGCTGCTGCT 1139
Qy 1336 ACCCAAGAGCCGGGACTCCCTTCTGCGACCCAGGCTGCTGCGGCGCCAGGCTGGGGT 1395
Db 1140 ACCCAAGAGCCGGGACTCCCTTCTGCGACCCAGGCTGCTGCGGCGCCAGGCTGGGGT 1199

1396 CCGCTCTCTCCACTCCAGGGCTCCGCCCCCAAGTGAAGGGGGCCCTGCGGAGCCTCAG 1455
1200 CGCTCTCTCCACTCCAGGGCTCCGCCCCCAAGTGAAGGGGGCCCTGCGGAGCCTCAG 1259
1456 ACACATGAGTTCAGGGCTGGGGGGCTTGGGCACATACCTGTCCTTGGCTATGACA 1515
1260 ACACATGAGTTCAGGGCTGGGGGGCTTGGGCACATACCTGTCCTTGGCTATGACA 1319
1516 GCTTTGGGGGCCCTTCCGGGAGCCAGCCCGGGGGCCGAGGTAGGTCTTGGGGGCTTAGAG 1575
1320 GCTTTGGGGGGCCCTTCCGGGAGCCAGCCCGGGGGCCGAGGTAGGTCTTGGGGGCTTAGAG 1379
1576 GCTTGGATGGCTCTGGGCCCAACCGCAGGGGGCAAGCGCAGGCCGGGCTGGAGGGCGGC 1635
1380 GCTTGGATGGCTCTGGGCCCAACCGCAGGGGGCAAGCGCAGGCCGGGCTGGAGGGCGGC 1439
1636 GCGGGGGCTCGGGCTGGGGGGCTCAGGTGAGCGCTCCGCTCCGGGGCTGGTCCGCGATCCC 1695
1440 GCGGGGGCTCGGGCTGGGGGGCTCAGGTGAGCGCTCCGCTCCGGGGCTGGTCCGCGATCCC 1499
1696 TCAGTCCCTCGGCCACCCGGGGGCTCGCTCCCTCGTCCCGACCGCACCCTGCGAGCCTCTT 1755
1500 TCAGTCCCTCGGCCACCCGGGGGCTCGCTCCCTCGTCCCGACCGCACCCTGCGAGCCTCTT 1559
1756 TGAACCGAGATCTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1815
1560 TGAACCGAGATCTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1619
1816 CATCTGATGGGTGGAATAATCACCGGGAATCCCCCTTCAGTTCTTTGAAAAAGTTCCA 1875
1620 CATCTGATGGGTGGAATAATCACCGGGAATCCCCCTTCAGTTCTTTGAAAAAGTTCCA 1679
1876 TGAACCGAGATCTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1935
1680 TGAACCGAGATCTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1739
1936 ATGCAATTTTAAATGGAATAATCAAACTGAAAGAACTGTTAGTGGCTTTAAAGCC 1995
1740 ATGCAATTTTAAATGGAATAATCAAACTGAAAGAACTGTTAGTGGCTTTAAAGCC 1799
1996 CCAAAAGCTCCCTAAGCGCTCCTCGAGATGAAGACCGGGGGAG - CCGCAGCCAGGTGGA 2054
1800 CCAAAAGCTCCCTAAGCGCTCCTCGAGATGAAGACCGGGGGAGCCCGCAGCGAGTGA 1859
2055 GACCCCGAGGACCGCGGGGGCCCGGTGACCGAGGCTCGCAGCGCGCCCGCTGAG 2114
1860 GACCCCGAGGACCGCGGGGGCCCGGTGACCGAGGCTCGCAGCGCGCCCGCTGAG 1919
2115 GGTCCGGCC - GAGCCAGGCTCCAGAGGGGGCGCGTTTGTCTCTCGGTTAAAAATAAGGTT 2173
1920 GGTCCGGCCGAGCGAGGCTCCAGAGGGGGCGCGTTTGTCTCTCGGTTAAAAATAAGGTT 1979
2174 CCGTCCG 2180
1980 CCGTCCG 1986

196 ATTTTCTCTGGGGCTCGGGGGCGCGGAGAGCTCATCCAGAGAGCGGTCCAGAGCG 255
1 ATTTTCTCTGGGGCTCGGGGGCGCGGAGAGCTCATCCAGAGAGCGGTCCAGAGCG 60
256 GACCCGGGAGTGTTCAGAGAGCGAGTGAAGAAGACAGGGGGCCCAAGTCCACAGCCAT 315
61 GACCCGGGAGTGTTCAGAGAGCGAGTGAAGAAGACAGGGGGCCCAAGTCCACAGCCAT 120
316 GCAGACCTGCGCCCTGGCATTCCTGCGCACGTTTCCAGGCGCTTGGGACCCCTCTCTT 375
121 GCAGACCTGCGCCCTGGCATTCCTGCGCACGTTTCCAGGCGCTTGGGACCCCTCTCTT 180
376 TTTGGCTGCTCTCTTGGTGTCTCAGAAATGAAGGCTGGGACAGCCCACTTGCACAGAGG 435
181 TTTGGCTGCTCTCTTGGTGTCTCAGAAATGAAGGCTGGGACAGCCCACTTGCACAGAGG 240
436 GGTAGTCTCTGTGTCTTGGGGCGAGAACCGGTCACTGCTGCAACATCTTCAATGAGGT 495
241 GGTAGTCTCTGTGTCTTGGGGCGAGAACCGGTCACTGCTGCAACATCTTCAATGAGGT 300
496 CTCCATGTCAACATCAAGCTGCTGCGCCAGCGGACAGAGCGCCATCTTCAATGAGGT 555
301 CTCCATGTCAACATCAAGCTGCTGCGCCAGCGGACAGAGCGCCATCTTCAATGAGGT 360
556 GGTCTCAGGCTACTTCTCCGGGACGCTGCGCTGCGAGTTCAGGGAGCGGTGGCACA 615
361 GGTCTCAGGCTACTTCTCCGGGACGCTGCGCTGCGAGTTCAGGGAGCGGTGGCACA 420
616 GGTCTGATCAAGAGCGCCCGGGACCTCCATGCTGGGGCTGACATGTGGGACCTCTGTTGG 675
421 GGTCTGATCAAGAGCGCCCGGGACCTCCATGCTGGGGCTGACATGTGGGACCTCTGTTGG 480
676 ACACAGAGAAATAACAGCAAGTCAAGTCAAGTGGAGTTCAGGTGAGAACCCAGTCCGC 735
481 ACACAGAGAAATAACAGCAAGTCAAGTCAAGTGGAGTTCAGGTGAGAACCCAGTCCGC 540
736 CCCTGACATGCGGTTCTGGCTGTCAGCGGTGCTCACTGCTGCTTCTTCTCTCTCT 795
541 CCCTGACATGCGGTTCTGGCTGTCAGCGGTGCTCACTGCTGCTTCTTCTCTCTCT 600
796 CGCTCTGGTCACTGTTCCGCTGTCAGAGTTCAGAGTTCAGAGTTCAGAGTTCAGAGTTCAGAG 855
601 CGCTCTGGTCACTGTTCCGCTGTCAGAGTTCAGAGTTCAGAGTTCAGAGTTCAGAGTTCAGAG 660
856 GTTCTTCTCTAGAACCCAGTGAAGTTCAGAGTTCAGAGTTCAGAGTTCAGAGTTCAGAGTTCAGAG 915
661 GTTCTTCTCTAGAACCCAGTGAAGTTCAGAGTTCAGAGTTCAGAGTTCAGAGTTCAGAGTTCAGAG 720
916 CCTGAGCAGAGCTCGCTGTAAGTTCAGAGTTCAGAGTTCAGAGTTCAGAGTTCAGAGTTCAGAG 975
721 CCTGAGCAGAGCTCGCTGTAAGTTCAGAGTTCAGAGTTCAGAGTTCAGAGTTCAGAGTTCAGAG 780
976 GGCACCTGGTTCACAAACCCCTCACTTGGAGCCCTTGGAGTTCGTCTCTCTCTCTCTCTCT 1035
781 GGCACCTGGTTCACAAACCCCTCACTTGGAGCCCTTGGAGTTCGTCTCTCTCTCTCTCTCT 839
1036 TGTTCATATGCGCAGACCCCATAGCCGCTGCAAGGCGAGAGGACACAGAGAGGACCA 1095
840 TGTTCATATGCGCAGACCCCATAGCCGCTGCAAGGCGAGAGGACACAGAGAGGACCA 899

Query Match 89.5%; Score 1951; DB 9; Length 2000;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1984; Conservative 0; Mismatches 0; Indels 3; Gaps 3;
US-09-997-165-3
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (119)..(865)
US-09-997-165-3

RESULT 3
US-09-997-165-3
; Sequence 3, Application US/09997165
; Patent No. US20020141999A1
; GENERAL INFORMATION:
; APPLICANT: Lyman, Stewart D.
; APPLICANT: Fanslow, William C.
; TITLE OF INVENTION: LIGAND FOR CD7 AND METHODS OF USE THEREOF
; FILE REFERENCE: 2913-US
; CURRENT APPLICATION NUMBER: US/09/997,165
; CURRENT FILING DATE: 2001-11-27
; PRIOR APPLICATION NUMBER: PCT/US00/14612
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/136,450
; PRIOR FILING DATE: 1999-05-28
; NUMBER OF SEQ ID NOS: 8

```
QY 1096 GGCCTGAGTGCAGACCTTGGGTGGCGGGGCTCTGTGTCACCACTCTCTGGTGGTGCAGAC 1155
Db 900 GGCCTGAGTGCAGACCTTGGGTGGCGGGGCTCTGTGTCACCACTCTCTGGTGGTGCAGAC 959
QY 1156 GACACGGCTTGTCTGGCAGGCTGGGCTCTGTGTCACCACTCTCTGGTGGTGCAGAC 1215
Db 960 GACACGGCTTGTCTGGCAGGCTGGGCTCTGTGTCACCACTCTCTGGTGGTGCAGAC 1019
QY 1216 CTTTCCCTCCACCCCGGAGCTTCCAAAGCTCTGTCTCTCAGTTTCCAAATGAACC 1275
Db 1020 CTTTCCCTCCACCCCGGAGCTTCCAAAGCTCTGTCTCTCAGTTTCCAAATGAACC 1079
QY 1276 ACTTACCTCCGAGACCCGACTTACAGAGAGGATGCCCTCTGCTCTGCTTATCAA 1335
Db 1080 ACTTACCTCCGAGACCCGACTTACAGAGAGGATGCCCTCTGCTCTGCTTATCAA 1139
QY 1336 ACCACAGACCGGACTCTCTTCTGTCACCCAGGCTGGTCCGGCCAGGCTGGGGT 1395
Db 1140 ACCACAGACCGGACTCTCTTCTGTCACCCAGGCTGGTCCGGCCAGGCTGGGGT 1199
QY 1396 CGGCTCTCTCCACTCCGAGGCTCCGCGCCCAAGTGAAGGGGCTCTGCGGAGCTCAG 1455
Db 1200 CGGCTCTCTCCACTCCGAGGCTCCGCGCCCAAGTGAAGGGGCTCTGCGGAGCTCAG 1259
QY 1456 ACACACTGAGTTTCAAGGCTGGGGGCTTGGCACATACCTGTCTCTTGGCTATGAGCA 1515
Db 1260 ACACACTGAGTTTCAAGGCTGGGGGCTTGGCACATACCTGTCTCTTGGCTATGAGCA 1319
QY 1516 GGCCTTGGGGGCTTCCGCGGAGCCCGGGGCGAGGTAGGTCTGGGGCTTAGAG 1575
Db 1320 GGCCTTGGGGGCTTCCGCGGAGCCCGGGGCGAGGTAGGTCTGGGGGCTTAGAG 1379
QY 1576 GCTGGATGGCTCTCTGGGCTCCAGCGGAGGCGCAAGCGAGCGGCTGGGAGCGCG 1635
Db 1380 GCTGGATGGCTCTCTGGGCTCCAGCGGAGGCGCAAGCGAGCGGCTGGGAGCGCG 1439
QY 1636 GCGCGGCTCGGGCTGGGGGCTCAGGTGGACGCTGCTTCCGCGGCTGGTCCGATCCC 1695
Db 1440 GCGCGGCTCGGGCTGGGGGCTCAGGTGGACGCTGCTTCCGCGGCTGGTCCGATCCC 1499
QY 1696 TCAGTCCCTCGGCGGCTCGGCTCTCTGTCACCTGTCGCGGAGGCTGCTTCTT 1755
Db 1500 TCAGTCCCTCGGCGGCTCGGCTCTCTGTCACCTGTCGCGGAGGCTGCTTCTT 1559
QY 1756 TGGACCCAGATCTTTTCTGCTTCTGTCACCTGTCGCGGCGGCTTGTGATGCTT 1815
Db 1560 TGGACCCAGATCTTTTCTGCTTCTGTCACCTGTCGCGGCGGCTTGTGATGCTT 1619
QY 1816 CATCTGTAAGGGTGAATAATCACCGGAATCCCTTCTTGTGAAAGTTCCA 1875
Db 1620 CATCTGTAAGGGTGAATAATCACCGGAATCCCTTCTTGTGAAAGTTCCA 1679
QY 1876 TGACTCGAATCTGAATGAAGAAACAAACCGACTCAAACTCCAGTAGTCCAA 1935
Db 1680 TGACTCGAATCTGAATGAAGAAACAAACCGACTCAAACTCCAGTAGTCCAA 1739
QY 1936 ATGCAATTTTAAATGGAATAACAAATCTCAAGAAACGCTTTAGTGGCTTAAAGC 1995
Db 1740 ATGCAATTTTAAATGGAATAACAAATCTCAAGAAACGCTTTAGTGGCTTAAAGC 1799
QY 1996 CCAAAACGTCCTTAAGGCTCTCGAGATGAAGACGGGGGGAG - CCCCAGCCAGGTGA 2054
Db 1800 CCAAAACGTCCTTAAGGCTCTCGAGATGAAGACGGGGGGAG - CCCCAGCCAGGTGA 1859
QY 2055 GACCCGAGAGACGGCGGCGGCTGAGACCGAGGCTTCGACAGCGGCGGCGGCTGAG 2114
Db 1860 GACCCGAGAGACGGCGGCGGCTGAGACCGAGGCTTCGACAGCGGCGGCGGCTGAG 1919
QY 2115 GGTCTGGGCT - GAGCCAGAGGCTCAAGAGGGGCGGCTTGTGCTCGGGTTAAATAGGTT 2173
Db 1920 GGTCTGGGCTGGAGCGGCTCAAGAGGGGCGGCTTGTGCTCGGGTTAAATAGGTT 1979
```

```
QY 2174 CCGTCCG 2180
Db 1980 CCGTCCG 1986

RESULT 4
US-10-775-169-97
; Sequence 97, Application US/10775169
; Publication No. US20040175743A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Burzynski, Michael
; APPLICANT: Twine, Natalie
; APPLICANT: Dorne, Andrew
; APPLICANT: Trepicchio, William
; TITLE OF INVENTION: Method for Monitoring Drug Activities In Vivo
; FILE REFERENCE: AM101080 (031896-013000)
; CURRENT APPLICATION NUMBER: US/10775,169
; CURRENT FILING DATE: 2004-02-11
; NUMBER OF SEQ ID NOS: 5278
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 97
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-775-169-97

Query Match 89.5%; Score 1951; DB 18; Length 2000;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1984; Conservative 0; Mismatches 0; Indels 3; Gaps 3;

QY 196 ATTTTCTCGGGCTCCGGGGCGGAGAAAGCTGCATCCAGAGGAGCGGTCCAGAGCG 255
Db 1 ATTTTCTCGGGCTCCGGGGCGGAGAAAGCTGCATCCAGAGGAGCGGTCCAGAGCG 60
QY 256 GACCCGGAGGTGTTTCAAGAGCAGTGACAGGACCAAGGCCCCAAGTCCCAAGCCAT 315
Db 61 GACCCGGAGGTGTTTCAAGAGCAGTGACAGGACCAAGGCCCCAAGTCCCAAGCCAT 120
QY 316 GCAGACCTGCCCCCTGGGCAATTCCTGGGCGGCTTCCAGAGGCTTGGAGCCCTCTGT 375
Db 121 GCAGACCTGCCCCCTGGGCAATTCCTGGGCGGCTTCCAGAGGCTTGGAGCCCTCTGT 180
QY 376 TTTGGCTGCTCTTGGTGTCTCAGAAATGAAGGCTGGGACAGCCCATCTGCACAGAGG 435
Db 181 TTTGGCTGCTCTTGGTGTCTCAGAAATGAAGGCTGGGACAGCCCATCTGCACAGAGG 240
QY 436 GGTAGTCTGTGTCTTGGGGGAGAACACCGTCTATGTCTGCAACATCTCCAAAGCCTT 495
Db 241 GGTAGTCTGTGTCTTGGGGGAGAACACCGTCTATGTCTGCAACATCTCCAAAGCCTT 300
QY 496 CTCCCATGTCAACATCAAGCTGCGTCCCGGAGGAGGAGCGCCATCTTCAATGAGT 555
Db 301 CTCCCATGTCAACATCAAGCTGCGTCCCGGAGGAGGAGCGCCATCTTCAATGAGT 360
QY 556 GGTCTCAGGCTACTTCTCCCGGAGCGGTGGGAGTCCAGGTTCCAGGAGGCGTGGCACA 615
Db 361 GGTCTCAGGCTACTTCTCCCGGAGCGGTGGGAGTCCAGGTTCCAGGAGGCGTGGCACA 420
QY 616 GCTGTGTATCAAGGCGCCGGGACTCCCATGTCTGGGTGTATGTGTGGACCTCTGTGG 675
Db 421 GCTGTGTATCAAGGCGCCGGGACTCCCATGTCTGGGTGTATGTGTGGACCTCTGTGG 480
QY 676 ACACAGAGAAATACAGCAAGTCAAGCTGGAGGTTTCAGGTGAGAACCCCAAGTCCG 735
Db 481 ACACAGAGAAATACAGCAAGTCAAGCTGGAGGTTTCAGGTGAGAACCCCAAGTCCG 540
QY 736 CCCTGACACTGGGTTCTGGGCTGTGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 795
Db 541 CCCTGACACTGGGTTCTGGGCTGTGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 600
QY 796 CGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 855
```

Db	601	CGCTCTGGTTCATGTTTCGCTGTCAGAGTGCCTGTTTCCAGCAACCGCGGAGAGAA	660
Qy	856	GTTCCTCTCTAGAAACCCAGATGAAGTTCGAGCCTCTAGAGCGGAGCCCGAGCAGG	915
Db	661	GTTCCTCTCTAGAAACCCAGATGAAGTTCGAGCCTCTAGAGCGGAGCCCGAGCAGG	720
Qy	916	CCTGAGCAGAGCTCCGCTGAACTCTGACCCCACTCCGAGCCCAACCCCAAGCCGCT	975
Db	721	CCTGAGCAGAGCTCCGCTGAACTCTGACCCCACTCCGAGCCCAACCCCAAGCCGCT	780
Qy	976	GGCACTGGTGTTCAAACCCCTCAACCTTGGAGCCTCGAGCTGCTGTCCTCCCAACCT	1035
Db	781	GGCACTGGTGTTCAAACCCCTCAACCTTGGAGCCTCGAGCTGCTGTCCTCCCAACCT	839
Qy	1036	TGTTTCCATATCCGAGACCCATAGCCGCTTGAAGGAGGAGGAGGAGGAGGAGGAGG	1095
Db	840	TGTTTCCATATCCGAGACCCATAGCCGCTTGAAGGAGGAGGAGGAGGAGGAGGAGG	899
Qy	1096	GGCCTGAGTCCGACCTTGGGTGGCGGCGCTGGGTCTCTGTCGCCACCCGAGGAGCACA	1155
Db	900	GGCCTGAGTCCGACCTTGGGTGGCGGCGCTGGGTCTCTGTCGCCACCCGAGGAGCACA	959
Qy	1156	GACACCGGCTTCTTGGCAGGCTGGGCTCTGTGTCAACCACTCTCTGGTGGTGCAGAC	1215
Db	960	GACACCGGCTTCTTGGCAGGCTGGGCTCTGTGTCAACCACTCTCTGGTGGTGCAGAC	1019
Qy	1216	CCTTCCCTCCACCCCGCAGGCTTCCAGCTCTGCTTCCAGTTCCTCAGTTTCCAAATGAA	1275
Db	1020	CCTTCCCTCCACCCCGCAGGCTTCCAGCTCTGCTTCCAGTTCCTCAGTTTCCAAATGAA	1079
Qy	1276	ACCTACCTCCGAGCACCAGCTTACAGAGGAGGATGCCCTCTCTGCCCTCATCAA	1335
Db	1080	ACCTACCTCCGAGCACCAGCTTACAGAGGAGGATGCCCTCTCTGCCCTCATCAA	1139
Qy	1336	ACCACAGACCCGAGCTCCCTTCTGCCACCCAGGCTGGTCCGCGCCAGGTGGGT	1395
Db	1140	ACCACAGACCCGAGCTCCCTTCTGCCACCCAGGCTGGTCCGCGCCAGGTGGGT	1199
Qy	1396	CGCTCTCTCCACTCCAGGCTCCGCGCCAGTGAAGGAGGAGGAGGAGGAGGAGGAG	1455
Db	1200	CGCTCTCTCCACTCCAGGCTCCGCGCCAGTGAAGGAGGAGGAGGAGGAGGAGGAG	1259
Qy	1456	ACACACTGGAGTTCAGGGCTGGGGGGCTTGGCAGATACCTGTCTCCCTTGGCTATGAGCA	1515
Db	1260	ACACACTGGAGTTCAGGGCTGGGGGGCTTGGCAGATACCTGTCTCCCTTGGCTATGAGCA	1319
Qy	1516	GGCTTTGGGGCCCTTCCGCGGAGCCCGCGGGCCGAGGTAGGCTGGGGCTTAGAG	1575
Db	1320	GGCTTTGGGGCCCTTCCGCGGAGCCCGCGGGCCGAGGTAGGCTGGGGCTTAGAG	1379
Qy	1576	GCTGGATGGCTCTGCGCCACCGCCAGGGGCAAGCGAGGCGGGCTGGGAGCGGC	1635
Db	1380	GCTGGATGGCTCTGCGCCACCGCCAGGGGCAAGCGAGGCGGGCTGGGAGCGGC	1439
Qy	1636	GGCGGGGCTCGGGCTGGGGGTCAGGTGGACGCTGCTCCCGGGCTGGTTCGCGATCCC	1695
Db	1440	GGCGGGGCTCGGGCTGGGGGTCAGGTGGACGCTGCTCCCGGGCTGGTTCGCGATCCC	1499
Qy	1696	TCAGTCCCTCGGCGACCCGGGGTGGCTCTGTCGCCACCGCACTGCGGAGCTCTT	1755
Db	1500	TCAGTCCCTCGGCGACCCGGGGTGGCTCTGTCGCCACCGCACTGCGGAGCTCTT	1559
Qy	1756	TGACCCAGATCTGTCATGCTTGTCTGTCACCTGGGGGGGCTTTCATGCTT	1815
Db	1560	TGACCCAGATCTGTCATGCTTGTCTGTCACCTGGGGGGGCTTTCATGCTT	1619
Qy	1816	CATCTGATGGGTGGAAAAATCACCGGAAATCCCTTCCAGTCTTTGAAAAAGTTCCA	1875
Db	1620	CATCTGATGGGTGGAAAAATCACCGGAAATCCCTTCCAGTCTTTGAAAAAGTTCCA	1679
Qy	1876	TGACTGAATATCTGAATGAAGAAACAAACCGACTCACAACCTCCAGTAGTCCCA	1935
Db	1680	TGACTGAATATCTGAATGAAGAAACAAACCGACTCACAACCTCCAGTAGTCCCA	1739

RESULT 5
US-09-799-777-109
; Sequence 109, Application US/09799777
; Patent No. US20020091244A1
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; Hillman, Jennifer L.
; Corley, Neil C.
; Guegler, Karl J.
; Baugh, Mariah
; Sather, Susan
; Shah, Purvi
; TITLE OF INVENTION: HUMAN SIGNAL PEPTIDE-CONTAINING PROTEINS
; NUMBER OF SEQUENCES: 154
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA: US/09/799,777
; FILING DATE: 06-Mar-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/002,485
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: BILLINGS, LUCY J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0459 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 109:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1064 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 109 :
US-09-799-777-109
Query Match 41.9%; Score 912.8; DB 9; Length 1064;
Best Local Similarity 96.1%; Pred. No. 1.4e-254;

```
Matches 957; Conservative 0; Mismatches 37; Indels 2; Gaps 2;
QY 199 TTCTCTGGGCTCCGGGCGCGAGAGCTGCATCCAGAGGAGCGGTCCAGAGCGGAC 258
Db 1 TTCTCTGGGCTCCGGGCGCGAGAGCTGCATCCAGAGGAGCGGTCCAGAGCGGAC 60
QY 259 CCGGGAGTGTTCAGAGCCAGTGACAAAGGACAGGGGCCCAAGTCCCAACGACCATGCA 318
Db 61 CCGGGAGTGTTCAGAGCCAGTGACAAAGGACAGGGGCCCAAGTCCCAACGACCATGCA 120
QY 319 GACCTGCCCTCCCTGGCAATCCCTGGCCACGCTTTCCAGGGCCCTTGGAGCCCTCTGTGTTT 378
Db 121 GACCTGCCCTCCCTGGCAATCCCTGGCCACGCTTTCCAGGGCCCTTGGAGCCCTCTGTGTTT 180
QY 379 GGCTGCTCTCTTGAAGTGTCAAGATGAAGGCTGGGACAGCCCATCTGACAGAGGGGT 438
Db 181 GGCTGCTCTCTTGAAGTGTCAAGATGAAGGCTGGGACAGCCCATCTGACAGAGGGGT 240
QY 439 AGTCTCTGTCTTGGGGGAGAACACCGTCTATGCTCTGCAACATCTCCAAAGCCCTTCTC 498
Db 241 AGTCTCTGTCTTGGGGGAGAACACCGTCTATGCTCTGCAACATCTCCAAAGCCCTTCTC 300
QY 499 CCATGTCAACATCAAGCTCGTGCCTCCAGCGGAGAGCGCATCTTCAATGAGTGGC 558
Db 301 CCATGTCAACATCAAGCTCGTGCCTCCAGCGGAGAGCGCATCTTCAATGAGTGGC 360
QY 559 TCCAGGCTACTTCTCCGGGACGGCTGGCAGTCCAGGTTTCAGGAGGGGTGGCAGGT 618
Db 361 TCCAGGCTACTTCTCCGGGACGGCTGGCAGTCCAGGTTTCAGGAGGGGTGGCAGGT 420
QY 619 GGTGATCAAGGGCCCGGACTCCCATGCTGGGCTGTACATGTGGACCTCTGGTGGACA 678
Db 421 GGTGATCAAGGGCCCGGACTCCCATGCTGGGCTGTACATGTGGACCTCTGGTGGACA 480
QY 679 CCAGAGAAATACAGACAGTCAAGTGGAGGTTTCAGGTGAGAACCCAGTCCGCGCCC 738
Db 481 CCAGAGAAATACAGACAGTCAAGTGGAGGTTTCAGGTGAGAACCCAGTCCGCGCCC 540
QY 739 TGACACTGGGTTCTGGCTGTGCCAGCGGTGTCTACTGCTGTTTCATCTCTTGGTGGC 798
Db 541 CGACACTGGGTTCTGGCTGTGCCAGCGGTGTCTACTGCTGTTTCATCTCTTGGTGGC 600
QY 799 TCTGTCATGTTGGCTGTGACAGTGGCTGTTTCCAGGAAACCCCGGGAGAAAGTT 858
Db 601 TCTGTCATGTTGGCTGTGACAGTGGCTGTTTCCAGGAAACCCCGGGAGAAAGTT 660
QY 859 CTTCCTCTAGAAACCCAGATCAAGTGGAGCGCTCAGAGCGGAGCCAGCAGGGCT 918
Db 661 CTTCCTCTAGAAACCCAGATCAAGTGGAGCGCTCAGAGCGGAGCCAGCAGGGCT 720
QY 919 GAGCAGAGCTCCGCTGAACTGTGACCCCGAGACTCCGAGGCCACCCCAAGGCCGCTGGC 978
Db 721 GAGCAGAGCTCCGCTGAACTGTGACCCCGAGACTCCGAGGCCACCCCAAGGCCGCTGGC 780
QY 979 ACTGTTGTTCAACCTCACCCTTGGAGCCCTGGAGCTGTGTCCTCCCGCCCAACCTTGT 1038
Db 781 ACTGTTGTTCAACCTCACCCTTGGAGCCCTGGAGCTGTGTCCTCCCGCCCAACCTTGT 839
QY 1039 TTCCATATGCGCAGACCATAGCGCTGCAAGGACAGAGGACACAGAGGAGCCAGCC 1098
Db 840 TTCCATATGCGCAGACCATAGCGCTGCAAGGACAGAGGACACAGAGGAGTACCAACC 899
QY 1099 CTGAGTCCGACCTTGGGTGGC-GGGGCTGGGTCTCTGTCCTCCACCGGAGGGCAGACA 1157
Db 900 CTGAGTCCGACCTTGGGTGGGCGGGCTGGGTCTCTGTCCTCCACCGGAGGGCACA 959
QY 1158 CACCGGCTTGTGGAGGCTGGGCTCTGTGTAC 1193
Db 960 AGACACCGGGCTTGTGGCAAGGCTGGGGCCCTC 995
```

RESULT 6
US-10-676-248B-245/c

```
; Sequence 245, Application US/10676248B
; Publication No. US20040161773A1
; GENERAL INFORMATION:
; APPLICANT: Rogan, Peter
; APPLICANT: Knoll, Joan
; TITLE OF INVENTION: SUBTELOMERIC DNA PROBES AND METHOD OF PRODUCING SAME
; FILE REFERENCE: 33026-B
; CURRENT APPLICATION NUMBER: US/10/676,248B
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: 60/415,345
; PRIOR FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: 10/676,248
; PRIOR FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: 60/494,494
; PRIOR FILING DATE: 2003-07-03
; NUMBER OF SEQ ID NOS: 251
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 245
; LENGTH: 3026
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; LOCATION: (1843)..(1843)
; OTHER INFORMATION: n is a, c, t or g
US-10-676-248B-245

Query Match 10.2%; Score 222; DB 18; Length 3026;
Best Local Similarity 95.4%; Pred. No. 1.8e-53;
Matches 250; Conservative 0; Mismatches 10; Indels 2; Gaps 2;

QY 2 TTCTGCTTCCTTAGCGTGAA-CGCGGGTGGGTGCTCCCGTGAATAATAAATTCAC 60
Db 1758 TTCTGCTTCCTTAGCGTGAAACGCGAGTGTGCTGCTCCGTGAATAATAAATTCAC 1699
QY 61 CGTCACGCTTGTGTGAACGCGGGTGTTCGCGAAACTTGGAGGCTTCCCGTAAACCCG 120
Db 1698 CGTCACGCTTGTGTGAACGCGGGTGTTCGCGAAACTTGGAGGCTTCCCGTAAACCCG 1639
QY 121 CTCCTTCCTCATCTGGAGGTGGTCCGCGCGGGTCCGCGCTCTCCCTGGCCCTTC 180
Db 1638 CTCCTTCCTCATCTGGAGGTGGTCCGCGCGGGTCCGCGCTCTCCCTGGCCCTTC 1580
QY 181 CCTCTGCTGCTTTTCATTTTCTGGGGTCCGCGGCGGAGAGCTGCATCCAGAGGA 240
Db 1579 CCTCTGCTGCTTTTCATTTTCTGGGGTCCGCGGCGGAGAGCTGCATCCAGAGGA 1520
QY 241 GCGGTCAGAGCGGACCCGG 262
Db 1519 GCGGTCAGAGCGGACCCGG 1498

RESULT 7
US-09-997-165-7
; Sequence 7, Application US/09997165
; Patent No. US20020141999A1
; GENERAL INFORMATION:
; APPLICANT: Lyman, Stewart D.
; APPLICANT: Fanslow, William C.
; TITLE OF INVENTION: LIGAND FOR CD7 AND METHODS OF USE THEREOF
; FILE REFERENCE: 2913-US
; CURRENT APPLICATION NUMBER: US/09/997,165
; CURRENT FILING DATE: 2001-11-27
; PRIOR APPLICATION NUMBER: PCT/US00/14612
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/136,450
; PRIOR FILING DATE: 1999-05-28
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 1112
; TYPE: DNA
; ORGANISM: Mus sp.
```

1807 TGATGCTTCATCTGATGGGTGGAAATACACCGGGAATCCCTTCAGTTCCTTTGAA 1866
1 TGATGCTTCATCTGATGGGTGGAAATACACCGGGAATCCCTTCAGTTCCTTTGAA 60

RESULT 9
US-10-450-826-83
; Sequence 83, Application US/10450826
; Publication No. US20040101818A1
; GENERAL INFORMATION:
; APPLICANT: Ji, Darren
; APPLICANT: Axelrod, Douglas W.
; APPLICANT: Cook, Jonathan S.
; APPLICANT: Jaiswal, Neelam
; APPLICANT: Eistein, Richard
; APPLICANT: Houghton, Adam
; APPLICANT: Mertz, Lawrence
; TITLE OF INVENTION: Gene Expression Profiles Associated with Osteoblast Differentiation
; FILE REFERENCE: 044921-5039-WO
; CURRENT APPLICATION NUMBER: US/10/450,826
; CURRENT FILING DATE: 2003-06-18
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: US 60/255,882
; PRIOR FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: US 60/285,691
; NUMBER OF SEQ ID NOS: 149
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 83
; LENGTH: 100301
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. AL020996
US-10-450-826-83

Query Match 2.5%; Score 54; DB 18; Length 100301;
Best Local Similarity 46.3%; Pred. No. 0.00052;
Matches 247; Conservative 0; Mismatches 285; Indels 2; Gaps 2;
QY 1189 GTCACCCACTCTGCGTGCAGACCCCTTCCCTCCACCCCGGAGCTTCCCAAGCTC 1248
DB 38881 GCCTTCCCCCATAGCCCGCCCGCAGGACCCCGCCCTCCGCTGCTCCCTCT 38940
QY 1249 TGTCTTCCTCAGTTTCCAAATGGAACACCTCACTCCGAGAGCAGCCAGTACCAAGAC 1308
DB 38941 CCGGTGCTCCCTCCGCTCCAGTCCCGCTCCG-CCTCGGCTCCCGGAGCTCGTCCATGC 38999
QY 1309 GCATGCCCTCCCTCTGCGCTCATCAACACCCAGACCCGAGTCCCTTTTGCACCC 1368
DB 39000 CCATGTCCCCCTGCTGCGGAGGCTGGGTCTTCACTCCCGCTCCCGGCTCCCGCC 39059
QY 1369 AGGCTGTCCGCGCCCGGAGGCTGCGGTCTCTTCACTCCAGGCTCCGCGCCCA 1428
DB 39060 ACCAGCTCGGCGCCCGGAGCTGCGCGCGGCTGCGGCTCCGCTGCGGCGGCGAGA 39119
QY 1429 GTGAGGGGGCCCTGCGGAGCTCAGACACTGAGTTCAGGCTCGGGGGGCTTGG 1488
DB 39120 CCGGAGCGGCGGCGGAGGCTGCGGCGGCGGAGGCTCCGAGAGCGCCAGGAGCGG 39179
QY 1489 CACATACCTGTCCCTTGGCTATGAGCAGGCTTTGGGGGGGCTTCCGCGGAGCCCGGG 1548
DB 39180 CCGGCGGCGGCTGCGGCGGAGGCTGCGCGGCGGCGGCGGCTGCGGCGGCGGCTT 39239
QY 1549 GCCGAGGTAGGCTTGGGGGCTTAGAGGCTGGGATGGTCTTGGCCCCACCGCCAG-GGG 1607
DB 39240 GCGCGGCGGCGGCGGAGGCTGCGCGGCGGCGGCGGCTGCGGCGGCGGCGGAGAGC 39299
QY 1608 GCAAGCGAGCGGCGGCTGCGGAGGCGGCGGCGGCTCGGCTCGGCTGCGGCGGCTCAGTGGAC 1667
DB 39300 GAAAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 39359
QY 1668 GCTGCTCCGCGGCTGCTGCGGAGTCCCTCAGTCCCTCGGCGGCGGCGGCGGCGGCTG 1721

FEATURE:
NAME/KEY: CDS
LOCATION: (135)..(773)
US-09-997-165-7

Query Match 7.1%; Score 155.4; DB 9; Length 1112;
Best Local Similarity 60.6%; Pred. No. 2.8e-34;
Matches 255; Conservative 0; Mismatches 166; Indels 0; Gaps 0;
QY 310 AGCCATGAGACCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCT 369
DB 131 AGCCATGAGACCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCT 190
QY 370 CCGTGTCTTGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCT 429
DB 191 CTTTCTACTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCT 250
QY 430 AGAGGGGGTAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 489
DB 251 TGAGCATGAAGTATCTGTAACAGAGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 310
QY 490 CGCCTTCTCCATGTCATCAATCAAGCTGGTGGTGGGAGAGACCGCTGCTGCTGCTGCT 549
DB 311 CAATCTCAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 370
QY 550 TGAGGTGGCTCAGGCTACTTCTCCGCGGAGCGGCTGCGGCTGCGGCTGCGGCTGCGGCT 609
DB 371 TCATAGCGCTCCAGGAACTACTCTAAGGATTCATGCGGAGCTTCAATATTCAGGAGTCCA 430
QY 610 GGCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 669
DB 431 GCGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 490
QY 670 CGTGGGACACACAGAGAAATAACAGACAGTCAAGCTGCGGCTGCGGCTTTCAGGTGCGAAGCCCA 729
DB 491 GCATGATTCAGGCGAGAGTTCAGAACTTCAACCTGACTGTTAATGCGGAGAGAGACA 550
QY 730 G 730
DB 551 G 551

RESULT 8
US-09-908-975-11235
; Sequence 11235, Application US/09089975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
; APPLICANT: SHOSHAN, Avi
; APPLICANT: WASSERMAN, Alon
; APPLICANT: MINTZ, Eli
; APPLICANT: FAIGLER, Simchon
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICING
; TITLE OF INVENTION: THAT POPULATE A TRANSCRIPTOME
; FILE REFERENCE: 36688-0005
; CURRENT APPLICATION NUMBER: US/09/908,975
; CURRENT FILING DATE: 2001-07-20
; PRIOR FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/287,724
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/221,607
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 32337
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 11235
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-908-975-11235

Query Match 2.8%; Score 60; DB 10; Length 60;
Best Local Similarity 100.0%; Pred. No. 5e-07;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 39360 AGGGGCTGGGGGGGGCTGGTTCCTCCGGCTTCGACGCCCGGCCGAGAGCCCG 39413

RESULT 10
US-10-723-860-3574
; Sequence 3574, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Gineburg, Wendy M.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; TITLE OF INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators
; CURRENT APPLICATION NUMBER: US/10/723,860
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 3574
; LENGTH: 100301
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-723-860-3574

Query Match 2.5%; Score 54; DB 18; Length 100301;
Best Local Similarity 46.3%; Pred. No. 0.00052;
Matches 247; Conservative 0; Mismatches 285; Indels 2; Gaps 2;

QY 1189 GTACACCACTCTGGGTGGTGCAGACCTTCCTCCACCCCGAGGCTTCGAGCTC 1248
Db 38881 GCCTTCCCGCCATCAGCCCGGCCAGGACCCCGCTCCCGCTCCCTCCTT 38940

QY 1249 TGCCTTCCAGTTTCCAAATGGAACCTCCTACCTCCGACACCCGACTTACAGGAC 1308
Db 38941 CCGGTGCTCCCTCCGTCAGTCCCTCCG-CTTCGCTCCCGGCTCCGACTCGTCAATGC 38999

QY 1309 GCATGCCCTCCCTCTGCTCCCTCATCAACCCAGACCCGAGCTCCCTTTCTGCCACCCC 1368
Db 39000 CCATGTCCCTCCCTGTGCGGAGGCTGGGTCTTTCAGTCCCGCTCCCGGCTCCCGCCC 39059

QY 1369 AGGCTGGTCCGCGCCAGGTGGGTGCGCTCTCTCACTCCAGAGGCTCCGCGCCCAA 1428
Db 39060 AGCCAGCTCGGCGCGGACCTGCGCGCGCGCTCGGCTCGGCTCGGCGGCGGCGAGA 39119

QY 1429 GTGAGGGGCCCCCTCCGAGGCTCAGACACTGAGTTCAAGGCTGGGGGGCTTGG 1488
Db 39120 CCGGACGCGCGGCGGCGGCGGCGGCGGCGGCTCCGAGCGCAGGGAACGGG 39179

QY 1489 CACATACCTGTCCCTTGGCTATGAGCAGGCTTTGGGGGGCTTCGCGCGGACGCCCGGG 1548
Db 39180 CGCGGGGCGGCTGGCGCGGAGGCTGCGCGCGGGGCGGGGCTGGCGGCGCGCTT 39239

QY 1549 GCCGAGGTAGGTCTTGGGGCTTAGAGGCTGGGATGGCTTCCTGGCCCAACCGCCAG-GGG 1607
Db 39240 GCGCGGCGCGGCGGCGGCGGCTGCGGCTGGCGGCGGCTGCGGCGGCGGCGGCGGAGC 39299

QY 1608 GCAAGCGGAGCGGCGGCTGGGAGGCGGCGGCGGCTCGGCTGGGGGCTCAGTGGAC 1667
Db 39300 GAAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 39359

QY 1668 GCTGCTCCGGGGCTGGTGGCGCATCCCTCAGTCCCTCGGCGCACCGGGGGTTCG 1721
Db 39360 AGGGGCTGGGGGGGGTGGTTCCTCCGGCTTCGACGCCCGGCCGAGAGCCCG 39413

RESULT 11
US-10-424-599-95310/c
; Sequence 95310, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J

; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 95310
; LENGTH: 606
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(606)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_57077C.1
US-10-424-599-95310

Query Match 2.4%; Score 52; DB 17; Length 606;
Best Local Similarity 50.2%; Pred. No. 0.00026;
Matches 154; Conservative 0; Mismatches 151; Indels 2; Gaps 1;

QY 1414 GGGCTTCGCGCCCAAGTAGAGGGGGCCCTCGCGGAGCCTCAGACACACTGGAGTTTCAGGG 1473
Db 567 GCGCGCCCGCCCGCCCGCGGGGGCGCGGGGGCGGCGCGGGGGCGGGGGCGGCGGCGG 508

QY 1474 CTGGGGGGGGCTTGGCACAATACCTGCTCTTGGGCTATAGCAGGCTTTGGGGGGCTTCC 1533
Db 507 GGGGGGGGGGGGGGGGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 448

QY 1534 GCGGCGAGCCCGGGGGCGAGTAGGCTCTGGGGGCTTAGAGGCTGGGATGCTCTCTGCG 1593
Db 447 GGGGGCGCGGG 388

QY 1594 CCCACCGCCAGGGGGCAAGCGAGCGCGGCTGGGAGCGG--CGGCGCGGCTCGGGCT 1651
Db 387 GGGGGGGGGGGGGGGGGGGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 328

QY 1652 GGGGGGTCAGTGGAGCTGCTCTCGGGGGCTGGTGGCGATCCCTCAGTCCCTCGGCCAC 1711
Db 327 GGGGGGGGGGGGGGGGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 268

QY 1712 CCGGGG 1718
Db 267 CCGGGG 261

RESULT 12
US-10-017-161-1481/c
; Sequence 1481, Application US/10017161
; Publication No. US20030143668A1
; GENERAL INFORMATION:
; APPLICANT: SUWA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 084335/0152
; CURRENT APPLICATION NUMBER: US/10/017,161
; CURRENT FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: JP 2001/246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2430
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 1481
; LENGTH: 5452
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; LOCATION: source

```

;
; FEATURE: (1)..(5452)
; LOCATION: (235)..(236)
; NAME/KEY: CDS
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE: (201)..(5252)
; LOCATION: (246)
; NAME/KEY: modified_base
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE: (1)..(140)
; NAME/KEY: modified_base
; LOCATION: (145)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE: (147)..(148)
; NAME/KEY: modified_base
; LOCATION: (162)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE: (165)
; NAME/KEY: modified_base
; LOCATION: (168)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE: (172)
; NAME/KEY: modified_base
; LOCATION: (189)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE: (191)..(193)
; NAME/KEY: modified_base
; LOCATION: (196)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE: (199)
; NAME/KEY: modified_base
; LOCATION: (214)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE: (218)
; NAME/KEY: modified_base
; LOCATION: (221)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE: (225)..(226)
; NAME/KEY: modified_base
; LOCATION: (229)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE: (232)
; NAME/KEY: modified_base
; LOCATION: (232)
; OTHER INFORMATION: a, t, c, g, unknown or other

```

```

;
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (235)..(236)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (246)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (248)..(252)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (256)..(257)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (261)..(263)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (265)..(266)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (271)..(273)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (276)..(277)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (283)..(285)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (296)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (298)..(300)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (302)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (307)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (314)..(315)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (334)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (339)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (343)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (345)..(348)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:

```

[illegible]

[illegible]

OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified base
LOCATION: (365)..(365)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified base
LOCATION: (367)..(368)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified base
LOCATION: (370)..(373)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified base
LOCATION: (376)..(378)
OTHER INFORMATION: a, t, c, g, unknown or other

Query Match 2.3%; Score 51.2; DB 17; Length 5452;
Best Local Similarity 20.0%; Pred. No. 0.0011;
Matches 194; Conservative 0; Mismatches 774; Indels 0; Gaps 0;
QY 459 AGAACACCGTCTGCTGCAACATCTCCACGCTTCTCCCATCTCAACATCAAGCTGC 518
DB 5270 AGGGNCTTCAATTGTTTAAANNCNNCTCTGCCCTCGNNNNNCCCTN 5211
QY 519 GTGCCACGGGAGGAGCGCATCTTCAATGAGTGGCTCCAGGCTACTTCTCCCGG 578
DB 5210 NNN 5151
QY 579 ACGGCTGGAGCTCCAGGTTACGAGGGGTGGCAGCTGTGTATCAAAAGCGCGCGG 638
DB 5150 NNN 5091
QY 639 ACTCCCATCTGGGTGTATGTGCGACCTCGTGGGACACAGAGAAATACACACAAG 698
DB 5090 NNN 5031
QY 699 TCACGCTGGAGGTTTCAGGTGCAGAACCCAGTCCGCGCTGACACTGGGTTCTGGCCTG 758
DB 5030 NNN 4971
QY 759 TGCCAGCGGTGTCTACTGTCTTCTATCTCTTGGTGGCTGTGTCTATGTTCTGGCTGGT 818
DB 4970 NNN 4911
QY 819 ACAGGTGCGCTGTTCCAGCAACCGCGGAGAGAGTCTTCTCTCTAGAACCCCGA 878
DB 4910 NNN 4851
QY 879 TGAAGGTGCGACGCTTCAGAGCGGAGCCAGCAGGCTGAGCAGAGCTCCGCTGAAC 938
DB 4850 CCCCCCCCCCNNNNNCCCCCCCCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCC 4791
QY 939 TGTGAGCCCCAGACTCCGAGCCCAACCAAGGCGCTGGCACTGTGTTCAAACCCCTCAC 998
DB 4790 CCCCCCCCCCNNNNNCCCCCCCCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCC 4731
QY 999 CACTTGGAGCCCTGGAGTGTGTCTCCCCCAGCCTTGTTCATATGCGCAGACCCA 1058
DB 4730 CNNNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCC 4671
QY 1059 TAGCGCTGCAAGGAGAGAGACACAGAGAGCGCAGCTGTAGTGGGCTGTTGGGTG 1118
DB 4670 CCCCCCCCCCNNNNNCCCCCCCCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCC 4611
QY 1119 GCGGGGCTGGGTCTCTCGTCCCAACCGAGGGGCAACAGACCCGCTTGTGGAGGCT 1178
DB 4610 NCCCCCCCCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCC 4551
QY 1179 GGGCTCTGTGTACCCACTCTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1238
DB 4550 NNNNNNNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCC 4491

QY 1239 TTCCAAGCTCTGCTTCTCAGTTTCCAAATGGAACCACTTCCCTCCGAGCACCAGAC 1298
DB 4490 CCCCCCCCCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCC 4431
QY 1299 TTACAGGAGCATGCGCTCTGCTCTCATCAAAACCCACAGACCCGAGCTCCCTTT 1358
DB 4430 NNNNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCC 4371
QY 1359 CTGCCACCCAGGCTGTGTCGCGCCCGCAGGTGTGGGTGCGCTCTCTCACTCCAGGGCT 1418
DB 4370 NNNNNNNNNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCC 4311
QY 1419 CCGCGCCC 1426
DB 4310 NCCCCCCC 4303
RESULT 14
US-10-424-599-61903
; Sequence 61903, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 61903
; LENGTH: 805
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(805)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_26913C.1
US-10-424-599-61903

Query Match 2.3%; Score 50.4; DB 17; Length 805;
Best Local Similarity 47.2%; Pred. No. 0.0086;
Matches 141; Conservative 0; Mismatches 158; Indels 0; Gaps 0;
QY 1420 CCGCCCAAGTGAGGGGCCCCCTGCGGAGCCTCAGACACACTGGAGTTTCAGGGCTGGGG 1479
DB 411 CCGGCGGGGTCCCCCGCNCNCGCGCCCGCGCGGGGGGGGGGGGGGGGGGGGGGG 470
QY 1480 GGGCTTGGCAATACCTGCTCTGCTGTATGAGCAGGCTTTGGGGGGCCCTTCGCGGCA 1539
DB 471 CCGGCGGG 530
QY 1540 GCGCCGG 1599
DB 531 GGG 590
QY 1600 GCCAGGGGGCAAGCGCAGCGCGGGCTGGGAGCGCGCGCGGGCTCGGGCTGGGGGGGTC 1659
DB 591 GGG 650
QY 1660 AGGTGGAGCTGCTCTCCGGGGGCTGTGTCGCGCATTCCTCAGTCCCTCGGGCACCCGGGGG 1718
DB 651 NGGG 709

RESULT 15
US-10-184-644-346/c
; Sequence 346, Application US/10184644

Sun Mar 20 14:17:44 2005

```
; Publication NO. US2003004930A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C227
; CURRENT APPLICATION NUMBER: US/10/184,644
; PRIORITY FILING DATE: 2002-06-28
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 346
; LENGTH: 671
; TYPE: PRT
; ORGANISM: Homo Sapien
; US-10-184-644-346

Query Match      2.3%; Score 50.2; DB 14; Length 671;
Best Local Similarity 6.8%; Pred. No. 0.00091;
Matches 39; Conservative 208; Mismatches 329; Indels 0; Gaps 0;

Qy 927 CTTCCGCTGAACCTGGACCCGACCTCCGAGCCACCCAGCGCGCTGGCACTGGTGT 986
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 648 H.H.C.YBHD.H.SCGY..D.MH.TS.CM.SAT.CMB.TS.HAS.MH.TM 589
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 987 TCRAACCTCACCATTGGAGCCTGGAGCTGTCTCCGCCCAACCTGTTCATAT 1046
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 588 ...T...CT...M.TM.M.KTNBM.BT..M.C.BYSM.YARB.ITM.KBHMNTMR 529
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1047 GCCGACAGCCATGACCGCTGTGACCCAGCCAGAGAGACAGAGAGCCCTGAGTGC 1106
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 528 YY.MM.ATBBHANM...SAB..AC....TN.GYMBH.SHB.M.TM..SD.M..BS. 469
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1107 CGACCTTGGTGGCGGCTGGTCTCTCTCCAGCTTCCAAATGGAACCACTCACCTCC 1166
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 468 .MM.BMYHS.C.SYVAY.BMB.MT..M..BYM.M.C.MYT..ASSS..MTSMWT 409
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1167 GCTTGGCAGGCTGGGCTCTGTGACCCAGCTCTGGGTGGTGCAGACCTTCCCTCC 1226
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 408 Y...T...SHSS..CYCYMYCYMB..HH.Y..HCSGCCCY.T.CYHTY.YYY.M. 349
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1227 ACCCCGAGGCTTCCAGCTCTCTCTCTCAGTTTCCAAATGGAACCACTCACCTCC 1286
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 348 .M...Y.YY.....YYTYT.HYTMW.S.YHB.HSHSHSSSS..Y..M.MYC 289
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1287 GCAGCACCGGCTTACCAAGCAGGATGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1346
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 288 Y.M...M.T.MYC..MMBSHSHSSSSSSSSSTSTYKTB...MTCHSHSHSHSTS. 229
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1347 CGGACTCCCTTCTCCACCCAGGCTGTGTCGGGCCAGGCTGGGTGGGTGGGTCTCTCT 1406
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 228 TMMMYCC..CYI.TYBTMM..A.H.HSAM.S.SSS..SN...S.SBST.H.HSSTMYT 169
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1407 ACTCCAGGCTCCGCGGAGTGGGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1466
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 168 MSBSKM.T.AMTM.CSNHSHSHSHSHS.KYHSTTATATATATATATATATATATAT 109
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1467 TTCAGGCTGGGGGCGCTTGGCATACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1502
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 108 T..T..SHSSSB...TSRSTD.NN...W.C.N.CMYM 73
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

Search completed: March 20, 2005, 09:46:02
Job time : 1262 secs